STATISTICAL AND CLINICAL APPROACHES TO THE IDENTIFICATION OF COUPLES AT RISK IN MARRIAGE

By

MARK BARRY LEFKOWITZ

A DISSERTATION PRESENTED TO THE GRADUATE COUNCIL OF
THE UNIVERSITY OF FLORIDA IN PARTIAL
FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA
1973

In Howie's Memory

You will always walk in the Kingdom of Mankind
And the Songbirds of Goodness
will spread their wings to share
Your smile forever.

ACKNOWLEDGMENTS

I first wish to express my sincere appreciation to Dr. Paul Satz, chairman of my doctoral committee, not only for his supervisory role in guiding my work on this dissertation, but for his interest and concern throughout my graduate career. I am privileged to know Dr. Satz as both mentor and friend and only hope that more would follow his example as a compassionate teacher and honest human being. His confidence in my work and his continual emotional support will always have special meaning to me.

Sincere thanks are also extended to Dr. Ben Barger, Dr. Gerald R. Leslie, Dr. William C. Mottola and Dr. Madelaine M. Ramey, members of my supervisory committee, for gently helping me grow as professional-scientist and ably assisting me in bringing this research to fruition.

Both Dr. Carl T. Clarke and Dr. Sidney M. Jourard are due my gratitude for serving as original members on my doctoral committee. I am especially grateful to Dr. Clarke for giving me access to his Marriage and College Life data bank and patiently awaiting my interpretations. He is an integral person behind this research although his name no longer appears on my committee.

Darryl Downing, my man behind the machine, deserves a great deal of thanks for his computer and statistical assistance. While Darryl may be a man of numbers and computer printouts, his genuine human qualities are too abundant to be measured.

A grateful note of thanks is also extended to those professional clinicians and students who participated as judges in this study.

I would also like to acknowledge the role that good friends played in helping me reach this goal. Their friendships were contingent on very little and their love was most strong during times of confusion and pain. Their beauty will always warm my soul wherever life takes me. They know who they are.

And to the folks at home I send a very special, silent message. One we have shared so many times before.

This research was supported in part by a grant from the National Institute of Mental Health (MH 15041).

TABLE OF CONTENTS

		Page
ACKNOWLEDGMENTS		iii
LIST OF TABLES		vii
LIST OF FIGURES		x
ABSTRACT		xi
CHAPTER		
I	INTRODUCTION	1
	General Overview	1
	Basic Purposes of the Study	3
	The Concept of Marital Adjustment	5
	Review of Actuarial-Clinical Literature	11
	Schematic for Presentation of Design and Data	16
II	THE STATISTICAL APPROACH: METHOD	19
	Subjects; Research Project Description	19
	Instruments	19
	The Independent Measures	23
	The Dependent Measure	29
	Procedure	38
	Conceptual and Terminological Issues	44
III	THE STATISTICAL APPROACH: RESULTS	47
	Analysis I: Standardization	47
	Analysis I: Cross-Validation	55
	Analysis II: Standardization	57
	Analysis II: Cross-Validation	67
	Comparison of the Two Analyses	70

TABLE OF CONTENTS (CONTINUED)

CHAPTER		Page
IV	THE CLINICAL APPROACH: METHOD.	73
	Rationale of Decision-Making Task. . . .	73
	The Marriage Profile Summary	74
	The Experimental Design.	75
	Summary of Procedure	88
V	THE CLINICAL APPROACH: RESULTS	90
	Basic Research Questions	90
	Hypotheses and Major Statistical Analyses.	91
	Importance of Didactic Training; Role of Experience	98
	Individual versus Group.	105
	Man versus Machine	111
	Other Findings	114
VI	DISCUSSION.	129
	The Statistical Model: Empirical Relationships in Marital Adjustment.	129
	The Clinical Decision-Making Task. . . .	138
	Comparison of Statistical and Clinical Approaches	145
	Community Prevention Proposal.	148
APPENDICES		
A	INSTRUMENTS ADMINISTERED TO MARRIED COUPLES .	155
B	EXAMPLES OF MARRIAGE PROFILE SUMMARY (MPS) FORM TAKEN FROM THE THREE ADJUSTMENT CATEGORIES.	184
C	TEST BOOKLETS USED BY JUDGES.	188
D	SUMMARY OF NEWMAN-KEULS TESTS	218
E	TYPOLOGY OF DECISION RULES GENERATED BY JUDGES.	221
REFERENCES.		234
BIOGRAPHICAL SKETCH		242

LIST OF TABLES

TABLE		Page
1	INDEPENDENT MEASURES DERIVED FROM MARRIAGE-RELATED QUESTIONNAIRES.	24
2	SUMMARY DATA FOR INTERQUARTILE DISTRIBUTION OF COUPLE MAS SCORES (N=235).	35
3	COMPARISON OF THE STANDARDIZATION AND CROSS-VALIDATION SAMPLES WITHIN CRITERION GROUPS.	42
4	DISCRIMINABLE RANK ORDER OF PREDICTOR VARIABLES IN ANALYSIS I	48
5	SCHEMATIC DESIGN FOR THE 2x2 CLASSIFICATION MATRIX.	49
6	CLASSIFICATION MATRICES FOR ANALYSIS I. . .	51
7	CROSS-VALIDATION MATRIX OF ANALYSIS I 10-VARIABLE MODEL.	56
8	DISCRIMINABLE RANK ORDER OF PREDICTOR VARIABLES IN ANALYSIS II.	58
9	SCHEMATIC DESIGN FOR THE 3x3 CLASSIFICATION MATRIX.	60
10	CLASSIFICATION MATRICES FOR ANALYSIS II . .	62
11	CROSS-VALIDATION MATRIX OF ANALYSIS II 5-VARIABLE MODEL.	68
12	CROSS-VALIDATION MATRIX OF ANALYSIS II 10-VARIABLE MODEL	69
13	UNIVARIATE RANK ORDER IMPORTANCE OF PREDICTOR VARIABLES TO CRITERION.	72

LIST OF TABLES (CONTINUED)

TABLE		Page
14	EXPERIMENTAL DESIGN SCHEMATIC	77
15	HYPOTHESES USED TO EVALUATE BASIC RESEARCH QUESTIONS	92
16	ANALYSIS OF VARIANCE OF THE MAIN TREATMENT EFFECTS AS MEASURED BY TOTAL HIT RATE (THR)	95
17	ANALYSIS OF VARIANCE OF THE MAIN TREATMENT EFFECTS AS MEASURED BY FALSE POSITIVE RATE (FPR).	96
18	ANALYSIS OF VARIANCE OF THE MAIN TREATMENT EFFECTS AS MEASURED BY FALSE NEGATIVE RATE (FNR).	96
19	ANALYSIS OF VARIANCE OF THE MAIN TREATMENT EFFECTS AS MEASURED BY MEAN CONFIDENCE SCORE (MCS)	97
20	ANALYSIS OF VARIANCE OF THE MAIN TREATMENT EFFECTS AS MEASURED BY APPROPRIATENESS SCORE (AS).	97
21	COMPARISON OF THE DIFFERENT TREATMENT GROUPS ON THE PRIMARY DEPENDENT VARIABLES .	99
22	ANALYSES OF VARIANCE OF THE MAIN TREATMENT EFFECTS ON THE GROUP DECISION-MAKING DATA .	106
23	MEAN SCORES FOR THE GROUP DECISION-MAKING TASK ON THE PRIMARY DEPENDENT VARIABLES . .	107
24	COMPARISON OF INDIVIDUAL AND GROUP ANOVAS FOR HOMOGENEITY OF VARIANCE	109
25	CELL MEAN SCORES AND t TEST DIFFERENCES FOR THE INDIVIDUAL AND GROUP DECISION-MAKING TASKS.	110

LIST OF TABLES (CONTINUED)

TABLE		Page
26	NUMBER AND TYPES OF DECISION RULES GENERATED BY JUDGES	118
27	SPEARMAN RANK-ORDER CORRELATIONS BETWEEN JUDGES AND COMPUTER MODEL	120
28	PEARSON PRODUCT-MOMENT CORRELATIONS BETWEEN JUDGES' WEIGHTS AND COMPUTER MODEL DERIVED WEIGHTS	122
29	ESTIMATES OF POPULATION BASE RATES MADE BY JUDGES FOR CRITERION GROUPS	124
30	HIT RATES AND DISTRIBUTION OF CLASSIFICATION FOR JUDGES EXACT IN THEIR ESTIMATES OF POPULATION BASE RATES	125
31	MEAN SELF-EVALUATION SCORES REFLECTING JUDGES' PRE- AND POST-TASK KNOWLEDGE.	127
32	QUALITATIVE IMPORTANCE OF INDEPENDENT VARIABLES USED IN PRESENT STUDY TO MARITAL ADJUSTMENT IN COLLEGE LIFE.	130

LIST OF FIGURES

Figure		Page
1	Judges' Mean Confidence Score (MCS) as a Function of Level of Experience.	102
2	Judges' Appropriateness Score (AS) as a Function of Experimental Conditions of Decision-Making Task	104
3	Diagrammatic Representation of Community Prevention Proposal.	153

Abstract of Dissertation Presented to the Graduate Council of the University of Florida in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy

STATISTICAL AND CLINICAL APPROACHES TO THE IDENTIFICATION OF COUPLES AT RISK IN MARRIAGE

By

Mark Barry Lefkowitz

March, 1973

Chairman: Paul Satz, Ph.D.
Major Department: Psychology

Both professionals and laymen have expressed noticeable interest in the human dynamics of marriage. Psychologists will be called upon to make many contributions to knowledge in this area. With the increasing emphasis on prevention programs and community mental health, psychologists would do well to address themselves to the identification of marital risk groups within specified populations. Intervention programs could then be developed to help those couples in trouble before problems became irreconcilable.

Given this perspective, the present research was designed to compare statistical and clinical approaches of identifying couples at risk in marriage. The university setting was chosen as an important community in which to gather data, limiting the investigation to college marrieds.

The study was divided into two separate stages of empirical investigation. The total design explored numerous research issues including: the psychology of married college

life; statistical classification models using multivariate analysis; actuarial versus clinical prediction; didactic applications of decision-making tasks; and prevention programs in community mental health.

The first phase was devoted to a step-wise discriminant analysis of data provided by 235 married couples. Twenty-eight predictor variables, psychological and demographic, were derived from responses to a battery of questionnaires. The empirical relationships between these variables and the adjustment criterion were appraised by partitioning the sample into three marital groups--At Risk, Borderline and Well-Adjusted. These criterion categories were determined by the couples' scores on a popular, well-validated marital adjustment test.

Several univariate and multivariate analyses were presented. The results generally showed that 18 of the predictors were significantly related to the adjustment criterion. The most important of these measures were (1) within-partner discrepancies between ideal role expectations and actual role enactments, especially in the areas of sexuality and solidarity, and (2) stress factors which monitor the amount of psychological discomfort experienced in the marriage-college environment.

A series of statistical classification models was generated by computer to assess actuarial skills in predicting marital adjustment. Multivariate prediction

equations were derived from a 10-variable computer model standardized on a sample of 175 couples. These equations correctly classified 73.1 percent of the cases. On a cross-validation sample of 60 couples, the computer hit rate dropped to 55.0 percent.

The second phase was addressed to a clinical judgment task. The objective was to evaluate how well judges could make predictions about married couples under conditions similar to those by which the statistical program operated. In addition to the man versus machine dimension several other questions were considered, including the importance of didactic training, the role of experience and the merits of group decision making.

The 36 judges, representing three levels of clinical-academic experience, were divided into treatment and control conditions. Both groups were asked to classify the 60 "marriage profiles" of the cross-validation sample. These profiles displayed the predictor data from the 10-variable computer model in graphic form. While judges in the treatment condition were allowed to study the profiles of the larger standardization sample (Didactic Training Session), the control judges were not provided such an opportunity.

The findings suggested that human judges performed no better or worse than the computer. No striking differences were found between treatment and control judges although there is some evidence that the didactic training promoted

distinct kinds of learning. The three experience groups did not differ in their accuracy although undergraduates expressed more confidence in their decisions. Some attitudinal differences were also noted. The group decision-making results were not consequential.

A research proposal is presented to demonstrate how the findings and methodology can be applied to a community prevention program for college marrieds.

CHAPTER I

INTRODUCTION

General Overview

In recent years clinical psychologists have taken a more genuine scientific and professional interest in the psychological dynamics of marriage. It is believed that in the not too distant future, psychologists--as researchers, as therapists, and as innovators of community prevention programs--will be called upon to make much greater contributions to our knowledge base in this area of human living.

Contemporary divorce trends suggest the need to focus on "marital adjustment" as a primary dimension of interest. A likely approach is to study the relationship of certain psychological and demographic variables to marital adjustment and to design a technique whereby these data can be used to maximally differentiate well- and poorly-adjusted marital couples.

The ultimate goal would be to enable the psychologist to incorporate his prediction techniques in the identification of couples who are "at risk" in marriage. A structured mental health intervention program might then be administered to ameliorate present adjustment difficulties and to minimize the chances of future marital breakdown.

At present, most clinical settings do not require the practicing psychologist to make decisions or statements about the "marital adjustment" of a man and woman. Rather, he is usually concerned with more traditional dependent variables--e.g., psychiatric diagnosis, selection of treatment modality, job success, neurological impairment, etc. It is not striking to note, therefore, that most psychologists would be unfamiliar with marital adjustment as a pertinent clinical criterion and would be less than comfortable in making predictions about this novel criterion using their traditional assessment tools.

The dissertation presented here is an attempt to elucidate some of the important parameters of marital adjustment in a limited population--college marrieds--while at the same time exploring different methodologies for facilitating learning in this area amongst mental health professionals. In actuality, this study encompasses a multifarious set of research issues: the psychology of married college life, statistical classification models using multivariate analysis, actuarial versus clinical prediction, decision-making tasks as a didactic training technique, and prevention programs in community mental health. This volume represents a compendium of multiple levels of data collection and the reader may be interested in its contents for one of various reasons. The more specific purposes of the dissertation are described below.

Basic Purposes of the Study

Four basic stages comprise the foundation of this research. These stages can be conceptualized as hierarchical levels of research since they are sequentially related to one another. That is, the latter stages can be developed experimentally only after the earlier stages have been carried out to completion. The four stages can be outlined as follows:

Analysis of Empirical Relationships

The first purpose of this study is to assess the empirical relationships between a number of psychological and demographic variables and the criterion of marital adjustment in a college population.

Development of a Computer Model

Given the above empirical information, the second purpose of the study is to create a statistical or computer classification model which can maximally differentiate particular levels of adjustment within this same population.

Evaluation of a Decision-Making Task

Using the data obtained from the computer classification model an individual decision-making task will be presented to persons of varying professional and academic experience. The purpose of this phase of the study is to determine whether such a task can be incorporated to "teach" clinical judges how to process novel test material in order

to make predictions about a novel criterion--marital adjustment. At the same time, the classification skills of the computer model will be matched against the classification skills of the clinical judges on a cross-validation sample of married couples.

A major premise of this research is that decision-making tasks such as the one designed for this study can be used in graduate training programs as excellent didactic models. Such a procedure allows the student to discover--through an active, in vivo learning experience--the empirical relationships between psychological and demographic variables and relevant clinical criteria.

It is believed that this phase of the research may show that individuals who have little or no knowledge about the psychology of marital adjustment can be quickly taught to develop a significant level of expertise in identifying couples "at risk." It may also be that clinical judges, using configural models of classification, can even surpass the accuracy of linear statistical models.

Conceptualization of a Campus Community Mental Health Program

With an ever-increasing emphasis on preventive approaches to problems of mental health (Cowen, Gardner & Zax, 1967) it is not unreasonable to assume that psychologists working in well-circumscribed communities--like the college campus--will soon be involved in designing pilot prevention programs for married persons.

The final purpose of this study is to demonstrate how the methodology and data of this research may be implemented in conceptualizing a secondary prevention program for married college students. Hopefully, it is possible to show how the statistical and clinical models can be used for selection purposes in the identification of an "at risk" group within the population. More importantly, by isolating the variables in these models which relate most significantly to the criterion, we can develop a more dynamic construct of "adjustment."

In this manner, by understanding more about the psychological "guts" of marital adjustment--what Holt (1970) would call a "job analysis" of the criterion--we are in a better position to design a viable prevention program whose structure is based on empirical findings.

The Concept of Marital Adjustment

Conceptualizing marital adjustment can lead to a good deal of theoretical and philosophical debate. It is a construct which is significantly less tangible than many other clinical outcome variables (for example, organic brain damage) and one which allows for great variability in attempting to reach a consensus about its dynamics. To further confound the issue, this criterion is as much a function of cultural values as it is a function of any so-called "socio-psychological reality."

Leslie (1967) provides a thorough discussion of this definitional problem as well as a review of the scientific attempts which have been made to systematically measure adjustment. The following is a brief overview of the approaches discussed in the literature which have been used to circumscribe the criterion.

Early work in this area was conducted on a fairly unsophisticated level (Davis, 1929; Hamilton, 1929; Hart & Shields, 1926). While the investigations of Bernard (1933, 1934) raised the quality of research considerably, the first classical study was conducted by Burgess and Cottrell (1939). Through the use of a self-report questionnaire the authors based their methodology on the following definition of adjustment:

> . . . a marriage in which the attitudes and acts of each of the partners produce an environment which is favorable to the functioning of the personality of each, particularly in the sphere of primary relationships. (Burgess & Cottrell, 1939, p. 10)

Given this perspective, the authors argued that adjustment could be measured by monitoring certain dimensions: (1) the amount of agreement between husband and wife on critical issues in the relationship; (2) the spouses' common interests and joint activities; (3) the overt demonstrations of affection and mutual confidences; (4) the number of complaints about the marriage; and, (5) feelings associated with loneliness, miserableness, irritability and lack of self-confidence.

They found that questionnaire items related to these dimensions correlated to a significant extent with marital happiness ratings supplied by a large sample of couples in the metropolitan Chicago area. In this sense, the composite adjustment score derived by Burgess and Cottrell served as a more systematic discriminant of success within a marriage than a crude happiness rating and helped provide a basic definitional foundation for the concept of marital adjustment.

A marital-happiness test constructed by Terman (1938) contained items oriented to similar dimensions as those emphasized by Burgess and Cottrell. However, Terman included other kinds of questions, as well. These dealt with the method of handling disagreements, the frequency of regretting marriage, the desire to marry the same partner again, the contemplation of divorce or separation, a general marital happiness rating, and an indication of the length of unhappiness.

A third kind of questionnaire was designed by Burgess and Wallin (1953) in their prediction study of adjustment in marriage from adjustment during engagement. This instrument was a more advanced version of the Burgess-Cottrell instrument and mitigated some of its limitations.

The systematic construction of a questionnaire was not the only approach used to define the adjustment criterion. Hart and Shields (1926) considered couples seeking marriage-license applications as adjusted while those appearing before

the Court of Domestic Relations were classified as maladjusted. Kirkpatrick (1937a, b, 1939) solicited ratings from the friends and relatives of married couples. These ratings were dichotomous judgments of either "well" or "poorly" adjusted.

Schroeder (1939) studied two samples of divorced and married couples. He classified the former as maladjusted and the latter as adjusted. Landis (1946) asked marrieds to rate themselves on a five-point scale in order to define adjustment. Less than 1 percent of his sample checked the lowest two categories ("unhappy" and "very unhappy")--which illustrates the weakness of this simplistic kind of procedure.

Most of the later research has continued to make use of the questionnaire technique in defining adjustment levels. The Burgess-Cottrell test has been used in numerous studies (King, 1952; Locke & Klausner, 1948; Locke & Mackeprange, 1949; Stroup, 1953; Williams, 1938). Terman's instrument or similar variants have been incorporated by Kelley (1939, 1941), Terman (1950) and Terman and Oden (1947). Also, Winch (1941) has reported data based on the Burgess-Wallin engagement-adjustment test.

The most popular contemporary approach to defining adjustment through the use of a numerical score stems from the work of Locke (1951). His major criticism of previous research was that construction of the adjustment questionnaires

was based on an unrepresentative sample of marrieds. The Burgess-Cottrell data were drawn from a sample consisting of 52.5 percent professionals, semi-professionals or upper-business types. Similarly, 56.5 percent of Terman's couples maintained professional, business-executive or managerial positions.

Locke devised a 29-item test which maximally differentiated between a sample of happily married persons and a sample of divorced persons--groups believed to represent the polarities of the adjustment continuum. Nineteen of the items were taken from the Burgess-Cottrell instrument, two were adaptations from Terman's items, and eight were formulated by the author himself. Locke reported that 72.5 percent of the happily married couples scored higher than the highest scoring divorced couple. Approximately three out of four of the divorced couples had scores less than the lowest scores from the happily married group. Cross-validation of this test on another population of marrieds is discussed by Locke and Karlsson (1952).

A shortened version of Locke's instrument has been reported in the literature (Locke & Wallace, 1959). This test was constructed by selecting those 15 items, modified when necessary, which had proved most discriminatory in previous questionnaire research. The test was administered to a pilot sample of 236 non-related spouses. Split half reliability was reported at r=.90.

From the larger sample of 236 Ss, two specific sub-samples of 48 Ss each were selected on which to cross-validate the questionnaire. The first sub-sample--a "maladjusted" group--consisted of 29 clients of the American Institute of Family Relations as well as two other persons for whom therapy case material was available, 11 recently divorced persons and six persons currently separated from their spouses. The second sub-sample--a "well-adjusted" group--consisted of only those persons judged to be exceptionally adjusted in marriage.

The differential performances by these two groups on the Locke-Wallace Questionnaire was obvious. The maladjusted sample obtained a mean score of 71.7 with only 17 percent of the total group scoring over 100. The well-adjusted sample obtained a mean score of 135.9 with 96 percent of the total group over 100.

In sum, the Locke-Wallace Marital Adjustment Test appears to be an accurate, reliable and parsimonious approach to measuring the criterion of interest in this study. It is probably the instrument most frequently reported in the current literature on marital satisfaction (e.g., Rogers et al., 1970) and is generally believed to be the best objective paper-and-pencil test available to researchers concerned with the dimension of adjustment. The following definition of marital adjustment, which can be inferred from the questionnaire, would appear appropriate and applicable in research with a college population:

> . . .the process of adaptation of the husband and wife in such a way as to avoid or resolve conflicts sufficiently so that the mates feel satisfied with the marriage and with each other, develop common interests and activities, and feel that the marriage is fulfilling their expectations. (Locke, 1951, p. 45)

Review of Actuarial-Clinical Literature

The actuarial-clinical controversy and its concommitant issues have been comprehensively outlined and documented by Meehl (1954) and other eminent psychologists (Chein, 1945; Gough, 1962; Holt, 1958, 1970; Lindzey, 1965; Sarbin, 1942; Towbin, 1960). The present research is an attempt to incorporate an innovative criterion--marital adjustment--in exploring some of the parameters of this important area of debate. It is by no means proposed as a definitive analogue of the controversy and makes no pretense of providing irrefutable answers to an area where research conclusions have almost uniformly been relative rather than absolute. For this reason an in-depth review of the literature will be avoided in favor of highlighting those historical issues which are pertinent to the data to be reported below.

Briefly stated, the controversy centers on defining the best--and in this sense, least cumbersome, most efficient, least time-consuming and most valid--methodology of predicting human behavior and making decisions regarding those clientele for whom psychologists are called upon to deliver mental health services. The problem is not confined only to the work of the clinician but also encompasses such similar

and overlapping fields of psychology as community, personality, social, abnormal, counseling, industrial, educational and vocational.

Actuarial or statistical prediction operates by a consideration of only statistical or numerical data. Individuals are classified into various statistical categories based on information drawn from previously studied individuals who have manifested similar antecedent traits, developmental histories, test data, etc. Attached to these categories are frequency rates or probability values related to the prediction criterion under consideration. The actuary uses these statistical relationships to determine the likelihood that the present individuals being examined will successfully meet or fail to meet the behavioral criterion.

In its purest form, the clinical or case study approach is an intuitive mode of prediction which incorporates both idiographic and nomethetic models. In this methodology a judge processes information about an individual using such data as overt behavior, clinical observations, and diagnostic test scores. This information is compared to the subjective and objective standards of human behavior which the judge brings to the decision-making task. His prediction about the individual is therefore based on this system of empirical rules, beliefs and philosophies that the judge has developed regarding the behavioral criterion. The clinician appraises the individual's relationship within this particular frame of reference in order to construct his prediction.

Therefore, in their most extreme forms, actuarial techniques depend upon statistical and mathematical tools while clinical techniques are oriented along intuitive, case study lines. Holt (1958) has challenged the appropriateness of this dichotomy and suggests that, while the clinician does not use well-defined formulae or equations in his judgment activities, he is still influenced by more than just qualitative data and often functions using covert kinds of statistical models. His differentiation between naive and sophisticated clinical techniques is an important one; its implications for this research will be returned to in a later section.

One of the initial documents cautioning against the wholesale use of actuarial approaches was published by Ray (1873) in his treatise on insanity. Early efforts in developing statistical tools of prediction frequently focused on issues of penology in determining criminal recidivism and parole success (Burgess, 1928; Glaser, 1955; Hart, 1923; Mannheim & Wilkins, 1955). This work was expanded by the Gluecks (1930, 1950, 1956) in their assessment of juvenile delinquency.

The utility of actuarial and clinical methods in vocational selection was first debated by Viteles (1925) and Freyd (1925). Shortly thereafter, Lundberg (1926) stressed the similarities of these methods instead of their differences. An avid proponent of the actuary, Sarbin (1942) devoted his early research to predicting achievement in college.

Much of the research in the actuarial-statistical controversy has been applied to the area of psychopathology and personality dynamics (Luborsky, 1962; Meehl, 1956, 1957; Phelan, 1964; Stricker, 1967; Wittman & Steinberg, 1944). Some of the other criteria which have been studied include: organic brain damage (Goldberg, 1959), length of hospitalization (Johnston & McNeal, 1967), interpersonal perception (Bradley & Stein, 1963; Cline & Richards, 1962), sporting event outcomes (Harris, 1963) and war game strategies (Merrifield & Erickson, 1965). There appear to be no empirical studies designed around marriage-related criteria. Also, the majority of the actuarial-clinical research has concerned itself with predicting to a dichotomous criterion. The subsequent study is an attempt to provide for a more complex and challenging decision-making task since the adjustment continuum will be defined in terms of three mutually exclusive and exhaustive criterion groups.

Research in this field has generated a number of correlative empirical issues which psychologists have isolated in their investigations. There are two basic variables within the clinical decision-making task itself which are also of heuristic relevance and to which a phase of this present research will be addressed. These independent variables are level of experience and individual versus group decision making.

Previous evidence on numerous kinds of decision-making tasks (Goldberg, 1959; Oskamp, 1962; Shagoury, 1969; Shagoury & Satz, 1969) has shown that as the amount of clinical experience increases, prediction accuracy decreases. In many of these studies judges were neither familiar with the criterion nor the methods used to predict to it. This inverse relationship between experience and accuracy did not hold in a study by Moxley and Satz (1970) where more experienced judges were more knowledgeable about the criterion and the predictor material provided them.

The role of experience will be examined in this study by determining the accuracy and appropriateness of judgments made by persons at three levels of expertise: Ph.D. clinical psychologists, graduate students pursuing a training program in clinical psychology, and undergraduate students taking courses in psychology. Although these judges will be dealing with a novel criterion and novel stimulus material it is hypothesized that the inverse relationship reported in previous research will not be demonstrated here. Since judges at all levels of expertise will be exposed to uniform training or teaching procedures it would appear to be theoretically justifiable to expect no differential effects. If it can be shown that undergraduates, graduates and professionals reap equal benefits from a didactic decision-making task, critics of traditionally lengthy professional training programs can add another piece of empirical evidence to support their claims that revisions are needed.

The previously mentioned Merrifield and Erickson (1965) research is one of the few studies to compare individual and group prediction. There is some theoretical logic to the argument that individual judges can maximize their predictive accuracy by merging together to make decisions using their collective hunches and insights about the data. Of course, the group process is subject to becoming unwieldy when the input made by its constituents is unmanageable and incapable of being synthesized. To monitor this question the experimental design of the present study will also consolidate individual judges into a group prediction session. It is hypothesized that the group decisions will have significantly higher hit rates than the individual decisions over all experience levels.

Schematic for Presentation of Design and Data

In some ways this volume is tantamount to two separate pieces of empirical investigation which, while complementary, have substance as independent areas of research. Because of the extensive nature of this study it has been decided to present the body of the material in a somewhat different and atypical format. Hopefully, this method of describing the experimental design and data will facilitate the reader's understanding of this project and will minimize the confusion which often results when numerous and intricate details must be reported.

Chapter II will describe the methodology used in the statistical phase of the study. A description of the data bank, subject pool and instruments administered will be given along with the independent measures derived from the various questionnaires. A major section of this chapter will be devoted to the operational procedures used to classify the dependent measure and the particular levels of marital adjustment. The statistical method of analysis will be outlined and the procedures of standardization and cross-validation explained. Finally, a few conceptual and terminological issues will be reviewed.

Chapter III will report the results obtained in the statistical phase of the study. Two kinds of discriminant analyses will be performed and standardization and cross-validation findings will be presented for both these analyses.

Chapter IV will be concerned with the methodology used in the clinical phase of the study. The rationale for the construction of the decision-making task will be presented and the paradigm for the experimental design systematically explained, including subjects, treatment and control conditions, independent variables, and dependent measures. The technique for presenting clinical information to judges through use of a Marriage Profile Summary will be detailed. Finally, the procedure for implementing the decision-making task will be discussed.

Chapter V will report the results obtained in the clinical phase of the study. These data will be reported in terms of four basic research questions and the major analyses performed to assess them. Ancillary test findings will also be provided.

Chapter VI will be devoted to a comprehensive discussion of the overall findings of the study. Data related to the empirical relationships in marital adjustment will first be appraised. Next, the clinical task will be evaluated and the data regarding role of experience and individual versus group decision making will be interpreted. Also, the didactic implications of this task will be explored. The third section of this chapter will compare the statistical and clinical approaches in terms of accuracy and utility. Finally, the relevance of these results will be amplified by outlining a brief proposal for a community prevention program in the campus setting.

CHAPTER II

THE STATISTICAL APPROACH: METHOD

Subjects; Research Project Description

The sample of subjects (Ss) consisted of 235 married couples who participated in various phases of the Marriage and College Life Project, a federally funded research investigation conducted at the University of Florida. The data were originally collected during the fall and winter quarters of the 1969-70 academic year, at which time at least one spouse in each participating couple was in attendance at the university. Those couples were selected for this phase of the research whose data files were complete, with no missing information pertinent to the statistical model to be described below. Future reference to Ss will be in terms of both individual data--i.e., information provided independently by the partners as separate individuals--and composite data--i.e., information provided mutually by the partners as a married unit.

Instruments

Psychological and demographic data believed relevant to the dynamics of married life were drawn from five questionnaires or self-report inventories which had been

administered to all Ss. These instruments will be described briefly and some can also be found in their entirety in Appendix A.

Marriage and College Environment Inventory (MACE)

This questionnaire determines those aspects of the college situation which are most commonly sources of stress or satisfaction to married college students. A stress factor can be derived for S which is a measure of the amount of psychological discomfort afforded by everyday situations in his or her living environment. Henceforth, for each couple tested, stress factor scores for both husband and wife are available. Appendix A contains the husband form of this 175-item questionnaire. A similar version of the MACE is given to the wife.

Kelly-Tharp Marriage Role Questionnaire

This 147-item questionnaire (Tharp, 1963a) contains two sections. Items in the first section assess those expectations which the partner brings into marriage--i.e., his values, beliefs and conventions. These expectations relate to certain roles of the husband and the wife as well as partnership roles. They monitor one's ideal notions about the manner in which husband and wife should conduct themselves in marriage. Items in the second section are analogous to those in the first but measure the actual or real ways in which husband and wife relate and behave in the marriage setting.

These role items can be classified into at least one of five factors or dimensions which relate to specific functions operable within marriage and family living. The five dimensions have been conceptualized (Tharp & Otis, 1966) as follows:

1. Solidarity. This dimension relates to the cohesiveness of a marriage. Its items monitor such elements as intimacy, social and emotional integration, togetherness, understanding and companionship.

2. Sexuality. This dimension is probably self-explanatory. Its items monitor such elements as premarital chastity, sexual fidelity and sexual gratification.

3. External Relations. This dimension relates to the couple's interaction with extrafamilial individuals and institutions. Its items monitor such elements as social activity and community affairs.

4. Internal Instrumentality. This dimension relates to the area of work-accomplishment within the household. Its items monitor such elements as wife's adequacy as housewife and husband's work performance.

5. Division of Responsibility. This final dimension relates to the authority and responsibility for originating, maintaining and terminating activities within the above-mentioned four areas. Its items monitor such elements as role-sharing, division of influences and masculine authority.

Appendix A contains the husband form of this instrument. A similar version is given to the wife.

Survey of Needs and Services for Married Students (Services Questionnaire)

This inventory is designed to collect descriptive biographical and demographic data on each couple and both partners respond to the questionnaire in cooperation with each other. The 97 items survey the following areas: housing, cultural-recreational activities, membership and organizations, counseling activities, employment, finances, medical expenses, family planning, day care services, and general background information.

Family Planning Evaluation Scale

This instrument measures the kinds of contraceptive techniques used by married couples and their attitudes and experiences regarding various family planning procedures. The husband and wife forms for this scale are virtually identical except the wife is asked to provide further information regarding children and related pregnancies.

Only one variable, number of children, was derived from this instrument and the entire scale is, therefore, not included in the appendix.

Locke-Wallace Marital Adjustment Test

The general nature of this 15-item questionnaire (see Appendix A) has been discussed in the previous chapter.

The Independent Measures

From Ss' responses to the first four questionnaires outlined above, 28 variables were isolated in order to assess their relationship to the marital adjustment criterion. Making up this set of independent measures are 19 scores which are psychological in nature and nine items which are demographic in nature. Table 1 presents these measures, their abbreviation codes when appropriate, and the questionnaires from which they were derived. They will be defined more explicitly below.

Stress Factor

This score, derived from S's responses to the MACE, is calculated by comparing those items assigned to the "generally stressful" category with those items assigned to the "generally satisfying" category. This ratio is then weighted by a percentage factor which allows for the number of items which are stress-neutral. The exact scoring procedure is described in Appendix A. The higher the Stress Factor score the greater the amount of psychological discomfort which S is experiencing in the marriage-college environment. Two independent measures are taken from this statistic: a Husband Stress Factor and a Wife Stress Factor.

TABLE 1

INDEPENDENT MEASURES DERIVED FROM MARRIAGE-RELATED QUESTIONNAIRES

Measure	Abbreviation Code	Questionnaire Derived From	Nature of Variable
Husband Stress Factor		MACE	
Wife Stress Factor			
Husband's Internal Instrumentality Score	H II		
Husband's Division of Responsibility Score	H DR		
Husband's Solidarity Score	H SL		
Husband's External Relations Score	H ER		
Husband's Sexuality Score	H SX		
Wife's Internal Instrumentality Score	W II		
Wife's Division of Responsibility Score	W DR	Kelly-Tharp Marriage Role Questionnaire	
Wife's Solidarity Score	W SL		
Wife's External Relations Score	W ER		Psychological
Wife's Sexuality Score	W SX		
Discrepancy in Attitude and Value Expectations/ Husband Items	DAVE H		
Discrepancy in Attitude and Value Expectations/ Wife Items	DAVE W		

TABLE 1 (CONTINUED)

Measure	Abbreviation Code	Questionnaire Derived From	Nature of Variable
Discrepancy in Attitude and Value Expectations/ Partnership Items	DAVE P		
Discrepancy in Perception of Role Enactments/Husband Items	DIPORE H		
Discrepancy in Perception of Role Enactments/Wife Items	DIPORE W	Kelly-Tharp Marriage Role Questionnaire	Psychological
Discrepancy in Perception of Role Enactments/Partnership Items	DIPORE P		
Realization of Companionship Score	R-C	Services Questionnaire	
Length of engagement			
Husband's age at marriage			
Wife's age at marriage			
Length of marriage			
Number of hours husband employed		Services Questionnaire	Demographic
Total annual income			
Husband's father's occupation			
Wife's father's occupation			
Number of children		Family Planning Evaluation Scale	

Tharp Discrepancy Scores

Various "role discrepancy" scores can be computed from the Kelly-Tharp Marriage Role Questionnaire by comparing each spouse's responses to the total battery of expectation and enactment items. Tharp (1963b) has concentrated on using intra-spouse discrepancy scores which are calculated to measure the differences between role expectations and role enactments within a particular partner. These scores are determined for each of the five role dimensions which have already been discussed. Ten independent measures are taken from this instrument. Five of these variables are tabulated to monitor husband expectation-enactment discrepancy in the following areas: Internal Instrumentality (H II score), Division of Responsibility (H DR score), Solidarity (H SL score), External Relations (H ER score) and Sexuality (H SX score). The remaining five variables are tabulated to monitor wife expectation-enactment discrepancy scores in the same areas (W II, W DR, W SL, W ER, and W SX scores). Higher scores on any of these variables indicate a greater amount of intra-spouse discrepancy.

DIPORE and DAVE Discrepancy Scores

Another mode of viewing the data provides an inter-spouse discrepancy parameter to the Kelly-Tharp questionnaire. The author, aided by other workers on the Marriage and College Life Project, designed two kinds of between-partner discrepancy scores which were computed for each couple.

The first measure monitors the differences in attitudes and values which husband and wife have individually brought into their marriage; the second measure monitors the differences in which husband and wife perceive roles actually being enacted in their marriage. The former will be mnemonically referred to as the DAVE (Discrepancy in Attitude and Value Expectations) variable while the latter will be known as the DIPORE (Discrepancy in Perception of Role Enactments) variable.

Six independent measures are taken from these statistics by partitioning the Kelly-Tharp items into husband-, wife- and partnership-related role categories. Three of these independent variables are computed to measure husband-wife discrepancy in their attitudes and values regarding roles exclusive to the husband (DAVE H), roles exclusive to the wife (DAVE W) and roles which are not spouse-specific but are rather partnership-oriented (DAVE P). The remaining three independent variables are computed to measure husband-wife discrepancy in their perception of the enactment of the above roles (DIPORE H, DIPORE W, and DIPORE P scores). For these measures, higher scores indicate a greater amount of inter-spouse role discrepancy.

Realization of Companionship Score

The final psychological variable is derived from two items on the Services Questionnaire which are analyzed using a multiplicative mathematical function. The Realization of

Companionship (R-C) score is a composite measure of a couple's need for companionship and the amount of time and energy they usually have for realizing this need. A couple's score can range from 1-18; the higher the score, the more likely a couple's need for companionship is great and being realized. The exact scoring procedure is described in Appendix A.

Demographic Variables

Nine background or demographic variables were chosen for the current study. The selection was based on findings previously reported in the literature (e.g., Burgess & Cottrell, 1939; Leslie, 1967; Terman, 1938) and on an extensive "eyeballing" of the Marriage and College Life Project data for apparent relationships with the adjustment criterion. It was also important that these data could be expressed in numerical terms. Eight of these independent measures were drawn from the Services Questionnaire: Length of engagement, Husband's age at marriage, Wife's age at marriage, Length of marriage; Number of hours husband employed, Total annual income, Husband's father's occupation, and Wife's father's occupation. The occupational variables were quantified on a five-point scale from professional (scored "1") to semi-skilled (scored "5"). Number of children was selected as the final demographic independent measure and was determined by analyzing information provided by female Ss on the Family Planning Evaluation Scale.

The Dependent Measure

The 28 independent variables which have now been outlined will be analyzed in terms of their relationship to marital adjustment--the dependent variable or criterion in the statistical model. The purpose of this section is to discuss the procedures used to operationalize the different levels of marital adjustment and to describe the manner in which Ss were classified along the dependent measure.

The Locke-Wallace Marital Adjustment Test was administered to all Ss. In its original form, a potential score can range from 2-158. Unfortunately, a significant number of Ss were given a questionnaire which was improperly composed and lacked one item (number 14). Consequently, all marital adjustment scale scores (MAS) reported in this research have been pro-rated based on a ceiling score of 143.

Two MAS scores were available for every couple in the sample--one each for the husband and wife. The dependent measure of interest is not, however, the adjustment level of an individual partner. Rather, we are concerned with the adjustment of the couple as a system of two persons. Therefore, it was necessary to combine the individual husband and wife MAS scores to obtain a measure of the couple's marital adjustment as a composite entity.

There appeared to be three possible methodological operations to choose from in determining this composite measure. The first approach assumes that a couple's general

adjustment is most reliably reflected by the more satisfied of the two members. Based on this approach, the larger of the two mates' MAS scores would be used to indicate the Couple MAS score. It is believed that this approach is a fallacious one--both theoretically and statistically--and it was not considered seriously.

The second approach assumes that a couple's general adjustment is most reliably reflected by the less satisfied of the two members. Based on this approach, the smaller of the two mates' MAS scores would be used to indicate the Couple MAS score. The premise of this approach is derived from a "weakest-link-in-the-chain" hypothesis and does carry a good deal of theoretical substance. However, it was felt that this method paid excessive emphasis to the independent role of a single spouse--rather than the interdependent or complementary roles of both spouses--in determining a couple's adjustment level. Accordingly, this technique was not adopted.

The third approach assumes that a couple's general adjustment is an amalgam of the degree of satisfaction experienced by each of the partners. Based on this approach, the mean of the two mates' MAS scores would be used to indicate the Couple MAS score. Locke (1951) feels that this combined score is justified and discusses composite data for his happily married and divorced couples.

There is varying evidence as to the expected amount of correspondence between husband and wife adjustment scores. Burgess and Cottrell (1939) report a spuriously high correlation of .88. Terman's (1938) correlational data were more realistic with an r=.59. Locke (1951) found some degree of correspondence between married spouses (r=.36) but none for divorced couples (r=.04). Other correlations which have been reported in the literature include: Harter (1950), r=.65; Burgess and Wallin (1953), r=.57 for engaged couples, r=.41 for married couples; Terman and Oden (1947), r=.52.

Given this perspective, mean MAS scores were computed separately for male and female Ss (husband (H) MAS = 102.0, SD = 19.1; wife (W) MAS = 105.4, SD = 18.0). These differences barely reached significance at the $p<.05$ level ($t = 1.99$). A Pearson product-moment correlation coefficient of .61 ($p<.001$) was obtained for these scores. In sum, the correspondence between the husband and wife adjustment scores in this study's sample seems to be as high or higher than most research findings reported in the literature.

The mean absolute difference between H and W MAS scores within couples was 12.9. From this information we can compare the "mean score" and "lower score" approaches to defining couple adjustment. That is, on the average, the Couple MAS score derived using the former approach is 6.45 units greater than the Couple MAS score derived using the latter approach. It therefore seems that there are no gross

statistical inconsistencies between these two techniques of defining couple adjustment and that the other findings discussed above support the use of the mean score procedure. For these reasons, this method was judged most suitable for the purposes of this study and it was adopted as the operation through which couple marital adjustment was defined.

The simplest division of the Couple MAS data would be into low and high adjustment groups using the mid-range score as the demarcation point. However, this breakdown of the data into dichotomous groups represents an arbitrary procedure based on no conceptual or methodological rationale.

The theory of psychological prophylaxis or prevention purports that psychological breakdown is a function of a compendium of stress conditions and that certain groups within a community can be identified as being more prone to these stressors. These groups are not conceptualized as already "sick" or "diseased"--to use the conventional medical metaphor. Rather, it is more probable that a person in one of these groups will contract the disease—entity in question than will a person in any other group within the total population. The concept of at risk has been popularized to describe this state of psychological liability or proneness.

This terminology is very consonant with the principles of selection theory (Meehl & Rosen, 1955). Different agencies sponsoring manpower programs--whether they be training programs for clinical graduate students, military programs for armed forces personnel, vocational programs for job

applicants, or therapy programs for hospitalized patients--are all concerned with the issue of selection. That is, they must choose a "treatment" group from a larger aggregate of persons that is most likely to profit from and maximize the impact of the particular program of interest. Every candidate is evaluated in relationship to the total pool of candidates in order to assess his "worthiness."

The prevention concept of at risk and the decision-making task of selection will serve as a conceptual foundation for classifying the MAS data as will a consideration of some important empirical findings.

Recent nationwide statistics (Lederer & Jackson, 1968) show that one out of four marriages eventually ends in divorce. Leslie (1967) has summarized three studies which have reported the proportion of happily married among the samples surveyed. Burgess and Cottrell (1939) found that 63.1 percent of 526 Illinois couples rated their marriages as very happy or happy; 14.4 percent reported average marriages, and 21.5 percent stated that their marriages were unhappy or very unhappy. Lang (1932) presented very similar statistics for 17,533 couples rated by friends: very happy or happy--61.6 percent; average--18.9 percent; unhappy or very unhappy--19.5 percent. Terman (1938) collected separate ratings for husbands and wives and reported a much higher proportion of satisfaction among 792 California couples. Terman's data show that, on the average, four out of five

spouses reported very happy or happy marriages and that only 5 percent considered their marriages at an unhappy or very unhappy level.

The Burgess-Cottrell, Lang and Terman data were gathered three to four decades ago when divorce rates were considerably lower. It is not presumptuous to suggest that different proportions of happiness and unhappiness would be reported using similar but contemporary samples of married couples. In fact, Burgess and Cottrell argue that the happiness ratings which they received reflect a very subjective dimension of measurement and are only partially indicative of the true nature of adjustment achieved in marriage. These authors divided their sample into three adjustment groups based on the objective scores obtained on their scale. Approximately equal proportions of the sample fell into "poor" and "fair" adjustment categories (28.5 percent and 28.3 percent, respectively). The rest of the sample (43.2 percent) was classified as achieving a "good" level of adjustment.

With the above conceptual-empirical perspective as a guide it was decided to generate a frequency distribution of the 235 Couple MAS scores. This distribution was divided into quartile segments and distinct levels of adjustment were operationalized using these groupings. Table 2 lists the summary data for the Couple MAS interquartile ranges.

The first quartile contains the lowest MAS scores in the sample. The mean MAS score for this group is 81.57 with a ceiling score of 93.5. All couples in this category fall

TABLE 2

SUMMARY DATA FOR INTERQUARTILE DISTRIBUTION
OF COUPLE MAS SCORES (N=235)

Quartile	Number of Cases	Range of Scores	Mean	Cumulative Frequency
1st	58	62.5--93.5	81.57	24.7%
2nd	60	94.0--104.5	99.00	50.2%
3rd	57	105.0--115.0	117.17	74.4%
4th	60	115.5--139.5		100.0%

considerably below the grand sample mean of 103.7. The 58 cases falling into this range of scores were therefore classified as an At Risk group. That is, in relationship to the other couples surveyed by the Marriage and College Life Project, these marriages appeared most vulnerable to breakdown and potential dissolution. For example, if a portion of the sample were to be chosen to receive the benefits of an intervention program, these couples would be considered to be "best bets" for such a selection.

It is interesting to note that the maladjusted group of marriages discussed by Locke and Wallace (1959) had a pro-rated mean MAS of 64.9. It is not contradictory for the At Risk Ss to have a higher adjustment mean than the comparable Locke-Wallace sample; in fact, this difference is to be expected. As discussed above, the maladjusted sample was at a recognizable level of marital despair. The concept of At Risk implies that a couple in this group has not reached such an obvious level of pathology but that the psychological momentum in the marriage is heading it in such a direction. Therefore, it seems appropriate--and operationally conservative--to refer to the lower quartile range of MAS scores as representing those couples who appear to be "at risk" in their marriages.

The second quartile contains the second lowest aggregate of MAS scores in the sample. The mean MAS score for this group is 99.00 with a ceiling of 104.5. Only seven

couples in this category fall above the grand sample mean. The 60 cases falling into this range of scores were thereby classified as a Borderline group. In general, these marriages are not as susceptible to breakdown as those identified as being at risk. However, their adjustment scores reflect signs and conditions which would classify them as less healthy than half the couples surveyed by the Marriage and College Life Project. In this sense, they occupy a "borderline" status in relationship to the total sample--demonstrating elements of both marital stability and duress.

The third and fourth quartiles contain the next-to-highest and highest segments of MAS scores in the sample. The mean MAS score for these quartiles taken as a composite group is 117.17 with a ceiling score of 139.5. All couples in this category fall above the grand sample mean. The 117 cases in this range of scores were thereby classified as a Well-Adjusted group. They differ only slightly on the adjustment continuum from the well-adjusted couples in the Locke-Wallace study who had an average pro-rated MAS of 123.0.

Through the marital adjustment questionnaire these couples report that they are coping well with the dynamics and demands of married life and are receiving considerable satisfaction from their partnerships. They tend to agree on many issues and when disagreement does arise the discord is resolved democratically and admirably. For these reasons, the rubric "well-adjusted" is a suitable label by which to refer

to this third group--both in consideration of the general validational findings which have been reported and in comparison with the other couples surveyed in the present sample.

Dunn's Multiple Comparisons of Group Means (Kirk, 1968) was run to assure that the At Risk (AR), Borderline (B) and Well-Adjusted (WA) groups were statistically different from each other. All three comparisons ($\overline{AR}$ vs. $\overline{B}$ = 17.43; $\overline{B}$ vs. $\overline{WA}$ = 18.17; $\overline{AR}$ vs. $\overline{WA}$ = 35.60) easily exceeded the critical Dunn value ($\underline{d}_{.01/2;3,232}$ = 4.70).

In summary, the dependent measure for this study has been operationalized in terms of a couple's marital adjustment score on the Locke-Wallace instrument. The 235 cases in the total sample have been partitioned into three levels of adjustment--At Risk, Borderline and Well-Adjusted. The procedure which was used to analyze the relationships between the 28 independent measures and these three criterion groups will now be set forth.

Procedure

Method of Analysis

The method of multivariate statistical analysis adopted for this research was the BioMed Step-Wise Discriminant Analysis Computer Program (BMD 07M) prepared at the University of California. The data were processed by the IBM 360 computer at the University of Florida Computer Center.

This step-wise analysis provides a large array of information. The technique is essentially a multiple regression problem except for the discontinuous distribution on the criterion variable. Some of its features are as follows:

1. Analyses of Variance. The program conducts one-way ANOVAs on each of the independent or predictor variables in relationship to the dependent or criterion variable.

2. Discriminable Variable Ranking (Step-Wise Regression). The program proceeds to select out, in a step-wise fashion, each predictor variable in terms of its discriminable relationship to the criterion groups which compose the dependent measure. This is done through covariance procedures. In this manner, an F to remove value is reported for each variable as it is selected by the program. The F value reflects the relative significance of that variable to the criterion after partialling out the variance accounted for by those variables which have already been removed in previous steps. In this manner, it is possible to rank order the independent variables in relationship to the dependent measure.

3. Development of a Computer Classification Model. At selected stages in the program lambda values are provided for those variables which have already been removed in the step-wise analysis. These lambda values usually differ for the same variable across the separate criterion groups. They are used as weights to derive discriminant function scores (Z scores) which maximally discriminate between the different

criterion groups. This is, in essence, what has been referred to in earlier sections as the "computer model of classification." These functions are used to classify each case (in this study, a married couple) into one of the criterion groups based on its predictor variable values and the lambda weights which have been assigned. The program also provides a classification matrix or contingency table at each of these stages. This matrix portrays how the model has classified the entire sample of cases into the different criterion groups. Hit rates and classification errors can be calculated using this matrix.

4. Classification of the Individual Cases. At the end of the program each case or married couple is evaluated in terms of its posterior probability of being classified into each of the criterion groups. These probabilities are based on the composite influence of all the variables fed into the program.

Standardization and Cross-Validation Samples

The adequacy of any classification model must be assessed by testing its predictive validity on a separate sample of cases (Holt, 1970; Satz, 1966). For this reason, the entire sample of 235 cases was not used to create the computer model. Instead, Ss were partitioned into two subdivisions: a standardization sample (N=175) which was used to create the classification model and a cross-validation sample (N=60) which was used to appraise the model's strength or accuracy.

This breakdown was accomplished by dividing each of the criterion groups--At Risk, Borderline and Well-Adjusted--into standardization and cross-validation samples. This process was conducted on a random basis within criterion groups with the only provision being that the ratio of cases in the criterion groups of the standardization sample be approximately equal to the ratio of cases in the criterion groups of the cross-validation sample (i.e., 1:1:2).

The MAS means for the standardization and cross-validation samples of each criterion group were compared using simple t-tests. No significant differences were found and the results are summarized in Table 3. These findings demonstrate that the standardization sample used to create the classification model and the cross-validation sample used to test the model are essentially comparable in terms of their adjustment criterion characteristics. Therefore, differences in this dependent variable cannot be held artifactually responsible for any hit-rate classification discrepancies between the standardization and cross-validation samples.

Two-Group versus Three-Group Discriminant Analyses

Two discriminant analyses were run in order to more thoroughly explore the relationship between the predictor variables and the adjustment criterion. The first program (Analysis I) was a two-group discriminant analysis designed to maximally discriminate between the At Risk and Well-

TABLE 3

COMPARISON OF THE STANDARDIZATION AND CROSS-VALIDATION SAMPLES WITHIN CRITERION GROUPS

	Standardization Sample			Cross-Validation Sample				
Criterion Group	N	M	SD	N	M	SD	df	t
At Risk	43	82.3	8.90	15	80.0	11.55	56	.797*
Borderline	45	99.2	3.37	15	98.5	4.19	58	.623*
Well-Adjusted	87	116.9	8.38	30	118.1	8.34	115	.696*

*not significant

Adjusted groups. It was hypothesized that this program would generate a highly accurate classification model since the two criterion groups represent the extremes of the adjustment continuum.

In this sense, Analysis I is somewhat academic in nature since it fails to take into consideration a substantial segment (25 percent) of the married population. Imagine a mental health program planned for college marrieds. It is unlikely that this program would select candidates or draw clients from a pool of marrieds who are exclusively either "at risk" or "well-adjusted." Also available in the pool of potential cases are marriages which we have classified as "borderline." Consequently, a computer model has real implications for a community prevention program only if it is derived from a discriminant analysis which exhaustively spans the adjustment continuum.

The second program (Analysis II) was a three-group discriminant analysis designed to maximally discriminate between the At Risk, Borderline and Well-Adjusted groups. It was hypothesized that this program would generate a classification model significantly better than chance (chance hit rate = 33.3 percent) but considerably less accurate than the Analysis I classification model. The three discrete criterion groups represent the entire adjustment dimension. This continuity along the dependent measure should obfuscate some of the more gross differences found when comparing only extreme groups.

Conceptual and Terminological Issues

To review, this phase of the study was designed to statistically appraise the relationship between the 28 independent or predictor variables (see Table 1) and the dependent measure of marital adjustment which has been operationalized into three mutually exclusive criterion groups. A few issues remain for discussion before proceeding to the results.

Prediction versus Classification

In a discussion of appropriate methodologies for validational studies Shneidman (1959) differentiates between three kinds of approaches to validity based on a temporal framework. Studies which are traditionally viewed as predictive are concerned with correlating the independent variables with future criteria. Similarly, a study is considered postdictive if it is concerned with correlating the independent variables with past criteria.

Shneidman suggests the term paridiction to define those studies which correlate their independent variables with more-or-less present criteria. This is what the APA Committee on Test Standards has referred to as status validity --the correlation between a test and concurrent external criteria.

Since the MAS criterion data were gathered from the married couples simultaneously with the other questionnaire data this study is not, in the strictest sense, exemplary of

prediction research. Rather, it is more classificatory in nature, since one set of data is being used to categorize persons on another set of data gathered at the same point in time. Consequently, what is here referred to as a predictor variable is in actuality a paridictor or classifier variable. To make research such as this predictive in the classical sense of the term we would need to gather information on couples early in marriage and then longitudinally assess their adjustment status through follow-up procedures.

Nonetheless, the terms prediction and classification are used interchangeably within the context of this research. By them is meant the task of incorporating a certain set of factors--operationally defined as the independent variable or predictor--in order to make statements about another set of factors--operationally defined as the dependent variable --irrespective of the temporal relationship between these two sets of factors. In short, these terms refer to the act of making statements about persons derived from stimulus material related to those same persons.

Sample and Population

The data obtained for this study have been sampled from married couples involved in academic-university life to one extent or another. In this sense, the empirical universe for this study consists of the total actual population of marrieds at the University of Florida while the conceptual universe consists of the total ideal population of all

married couples involved academically with university communities. Consequently, any attempt to extrapolate from these findings to all marrieds across the entire population would be scientifically premature and suspect. This research deals with a very select population of college marrieds and any generalizations beyond this population must be made with extreme caution.

CHAPTER III

THE STATISTICAL APPROACH: RESULTS

Analysis I: Standardization

The purpose of this discriminant function analysis was to create a classification model using the 28 predictor variables for two discrete criterion groups--At Risk (N=43) and Well-Adjusted (N=87). Table 4 lists the predictor variables in their order of discriminable ranking as determined by the step-wise analysis.

The F to Remove statistic indicates the F value for each variable as it was selected by the program in a step-wise fashion. This value monitors the relative significance of the predictor to the criterion in relationship to those predictors already selected. The Independent F statistic indicates the F value for each variable as determined by a one-way ANOVA. This value monitors the univariate significance of the predictor to the criterion independent of the other predictors in the program.

The program was written to generate five classification matrices at different stages of the step-wise analysis based on the lambda weights for the relevant predictor variables at each stage. Table 5 provides a schematic design of how

TABLE 4

DISCRIMINABLE RANK ORDER OF PREDICTOR VARIABLES IN ANALYSIS I

Predictor Variable	Rank	F to Remove	Independent F
Husband Stress Factor	1	63.23***	63.23***
H SX	2	33.92***	60.54***
W SL	3	16.39***	48.05***
H DR	4	6.81***	25.74***
DAVE H	5	4.68*	2.82
Wife Stress Factor	6	3.46	29.82***
Husband's age at marriage	7	3.59	6.29**
R-C	8	3.51	20.36***
Length of engagement	9	2.31	3.58
Number of hours husband employed	10	3.17	3.06
H ER	11	2.29	23.50***
DIPORE H	12	1.35	10.53***
Husband's father's occupation	13	1.18	0.01
W ER	14	1.26	9.55***
H SL	15	0.73	46.19***
H II	16	0.78	27.97***
DAVE W	17	0.62	1.62
DIPORE P	18	0.67	18.12***
Length of marriage	19	0.59	0.23
W DR	20	0.58	6.24**
DAVE P	21	0.48	5.83**
W II	22	0.18	4.92*
DIPORE W	23	0.24	8.73***
Wife's age at marriage	24	0.20	6.45**
Wife's father's occupation	25	0.22	1.17
Number of children	26	0.14	2.42
W SX	27	0.09	33.50***
Total annual income	28	0.00	0.28

*$p<.05$
**$p<.025$
***$p<.001$

TABLE 5

SCHEMATIC DESIGN FOR THE 2x2 CLASSIFICATION MATRIX

Classification by Function	Criterion Classification		Total N Classified by Function
	At Risk	Well-Adjusted	
At Risk	Cell 1 Valid Positives	Cell 2 False Positives	Cells 1 + 2
Well-Adjusted	Cell 3 False Negatives	Cell 4 Valid Negatives	Cells 3 + 4
Total in Class	Cells 1+3	Cells 2+4	Cells 1+2+3+4

to read a 2x2 matrix. The column categories refer to the actual adjustment status of the cases as measured by the criterion while the row categories refer to how the cases have been classified by the discriminant functions (which we can loosely refer to as a "test").

Given this outline, it is possible to identify what each cell in the matrix represents. All those cases falling into Cell 1 are called valid positives. These are married couples in the At Risk sample who have been accurately classified by the test as being at risk. All those cases falling into Cell 2 are called false positives. These are married couples in the Well-Adjusted sample who have been incorrectly classified by the test as being at risk. Of the two possible classification errors, the false positive error is the less serious in terms of this study.

All those cases falling into Cell 3 are called false negatives. These are married couples in the At Risk sample who have been incorrectly classified by the test as being well-adjusted. In terms of the implications and applications of this study, the false negative error is considered to be a more serious diagnostic mistake than a false positive error. Finally, all those cases falling into Cell 4 are called the valid negatives. These are married couples in the Well-Adjusted sample who have been accurately classified by the test as being well-adjusted.

The classification matrices for Analysis I are given in Table 6. Matrix A was generated at Step 5 after the program

TABLE 6

CLASSIFICATION MATRICES FOR ANALYSIS I

MATRIX A: 5-VARIABLE COMPUTER MODEL

Classification by Function	Criterion Classification		Total N Classified by Function
	At Risk	Well-Adjusted	
At Risk	34	5	39
Well-Adjusted	9	82	91
Total in Class	43	87	130

Hit Rate = 89.2%
False Positive Rate = 5.7%
False Negative Rate = 20.9%

MATRIX B: 10-VARIABLE COMPUTER MODEL

Classification by Function	Criterion Classification		Total N Classified by Function
	At Risk	Well-Adjusted	
At Risk	36	5	41
Well-Adjusted	7	82	89
Total in Class	43	87	130

Hit Rate = 90.8%
False Positive Rate = 5.7%
False Negative Rate = 16.3%

TABLE 6 (CONTINUED)

MATRIX C: 15-VARIABLE COMPUTER MODEL

Classification by Function	Criterion Classification		Total N Classified by Function
	At Risk	Well-Adjusted	
At Risk	36	4	40
Well-Adjusted	7	83	90
Total in Class	43	87	130

Hit Rate = 91.5%
False Positive Rate = 4.6%
False Negative Rate = 16.3%

MATRIX D: 20-VARIABLE COMPUTER MODEL

Classification by Function	Criterion Classification		Total N Classified by Function
	At Risk	Well-Adjusted	
At Risk	36	5	41
Well-Adjusted	7	82	89
Total in Class	43	87	130

Hit Rate = 90.7%
False Positive Rate = 5.7%
False Negative Rate = 16.3%

TABLE 6 (CONTINUED)

MATRIX E: 28-VARIABLE COMPUTER MODEL

Classification by Function	Criterion Classification		Total N Classified by Function
	At Risk	Well-Adjusted	
At Risk	36	5	41
Well-Adjusted	7	82	89
Total in Class	43	87	130

Hit Rate = 90.7%
False Positive Rate = 5.7%
False Negative Rate = 16.3%

had selected the five most discriminable variables. In a similar manner, Matrix B was generated at Step 10, Matrix C at Step 15, Matrix D at Step 20 and Matrix E at Step 28. Hit rates and the percentage of false positive and false negative classifications are given for each matrix.

From a univariate perspective, the results of Analysis I show that 19 of the predictor variables are significantly related to the criterion. The At Risk married couples showed higher scores on all 10 Tharp discrepancy measures, DAVE P, all three DIPORE scores, and both Husband and Wife Stress Factors. These same couples showed lower scores on R-C, Husband's age at marriage and Wife's age at marriage. Only two of the 19 significant variables were demographic in nature.

From a multivariate perspective, only the first five variables selected by the step-wise analysis account for significant variance in the criterion groups. These five variables are Husband Stress Factor, H SX, W SL, H DR and DAVE H. Combining these five variables into a multivariate classification model results in an extremely high hit rate of 89.2 percent (see Table 6, Matrix A).

If the 10-variable classification model (Table 6, Matrix B) is used the hit rate is improved (90.8 percent) but not considerably. A 10-variable model may seem to be an inefficient and costly procedure when a 5-variable model will predict almost as well. Nonetheless, this argument

fails to take into consideration the fact that two fewer false negative misclassifications are made by the larger computer model. Since the false negative error is deemed a serious one the results indicate that the 10-variable classification model may have merits over the 5-variable one.

The 15-variable model (Table 6, Matrix C) elevates the hit rate only slightly (91.5 percent) but adds no further improvement to the false negative rate. The 20- and 28-variable models revert back to the hit rate of the 10-variable model with similar false negative and false positive rates. In essence, these three models appear relatively inefficient compared with the former two.

Given the above findings, it was decided to cross-validate the classification program using the 10-variable computer model.

Analysis I: Cross-Validation

The purpose of this analysis was to cross-validate the 10-variable classification model on a separate sample of At Risk and Well-Adjusted couples. This was accomplished by applying the discriminant functions of this model to the appropriate data from the remaining At Risk (N=15) and Well-Adjusted (N=30) cases.

Table 7 shows the classification matrix for this cross-validation procedure. Inspection of the table shows that the 10-variable model correctly classified 84.4 percent of

TABLE 7

CROSS-VALIDATION MATRIX OF ANALYSIS I
10-VARIABLE MODEL

Classification by Function	Criterion Classification		Total N Classified by Function
	At Risk	Well-Adjusted	
At Risk	12	4	16
Well-Adjusted	3	26	29
Total in Class	15	30	45

Hit Rate = 84.4%
False Positive Rate = 13.3%
False Negative Rate = 20.0%

the total sample--a shrinkage of only 6.4 percent in overall hits compared to the original standardization. Also, there was a slight, but by no means drastic, increase of 3.7 percent in the false negative rate.

Based on the standardization and cross-validation findings for Analysis I it can be concluded that the 10-variable classification model is a powerful device for discriminating between At Risk and Well-Adjusted criterion groups. The hit rates are extremely high and it would be difficult for a clinical judge to match them. This appears to be an excellent case where actuarial procedures are most appropriate for the task of decision making.

The next analyses considered the three-group criterion problem in an attempt to assess how statistical techniques would fare under these conditions.

Analysis II: Standardization

The purpose of this discriminant function analysis was to create a classification model using the 28 predictor variables for three discrete criterion groups--At Risk (N=43), Borderline (N=45) and Well-Adjusted (N=87). Table 8 lists the predictor variables in their order of discriminable ranking as determined by the step-wise analysis. Also included are the F to Remove and Independent F values for each predictor.

As in Analysis I this program generated five classification matrices at Steps 5, 10, 15, 20 and 28. However,

TABLE 8

DISCRIMINABLE RANK ORDER OR PREDICTOR VARIABLES IN ANALYSIS II

Predictor Variable	Rank	F to Remove	Independent F
H SX	1	33.96***	33.96***
Husband Stress Factor	2	18.48***	30.05***
Wife Stress Factor	3	10.37***	17.93***
R-C	4	5.93***	15.07***
W SL	5	3.81**	24.31***
H DR	6	2.78	14.39***
Number of children	7	2.30	3.19*
Wife's father's occupation	8	2.24	2.25
Wife's age at marriage	9	1.76	3.99**
Husband's age at marriage	10	1.94	3.05*
Length of engagement	11	1.98	1.78
Number of hours husband employed	12	2.07	1.52
DIPORE P	13	2.08	10.81***
DAVE H	14	1.60	1.86
H ER	15	1.36	11.91***
DIPORE H	16	1.31	6.90***
H II	17	1.02	14.42***
H SL	18	1.76	26.88***
Total annual income	19	0.90	1.12
W II	20	0.49	2.43
DIPORE W	21	0.50	4.52**
DAVE W	22	0.66	1.35
W SX	23	0.27	16.00***
W DR	24	0.21	3.07*
W ER	25	0.17	4.24**
Husband's father's occupation	26	0.13	0.12
DAVE P	27	0.10	2.66
Length of marriage	28	0.03	0.51

* $p<.05$
** $p<.025$
*** $p<.001$

since three criterion groups are involved, the matrices are somewhat more complex. Table 9 provides a schematic design of how to read a 3x3 matrix. The manner of identifying what each cell represents does not change. The terminology used to categorize classification hits and errors differs slightly. The operations for defining false positive and false negative errors have been developed by the author.

Cells 1, 5 and 9 contain those cases which have been validly classified. These are married couples in the At Risk, Borderline and Well-Adjusted samples who have been accurately classified by the test into their respective adjustment categories. The hit rate can be computed by adding the number of cases in these three cells along the left-to-right diagonal.

The rest of the cells represent misclassified cases. The following logic is used to identify which cells contain false positive and false negative errors. A false positive error is defined as when a test classifies a case as more "pathological" than it is in actuality. In terms of the criterion of interest here, a false positive error is defined as when the test classifies a married couple as being less adjusted than it is in actuality.

Cells 2, 3 and 6 represent false positive classifications. Cases in Cell 2 are married couples in the Borderline sample who have been incorrectly classified by the test as being at risk. Cell 3 cases are married couples in the

TABLE 9

SCHEMATIC DESIGN FOR THE 3x3 CLASSIFICATION MATRIX

Classification by Function	Criterion Classification			Total N Classified by Function
	At Risk	Borderline	Well-Adjusted	
At Risk	Cell 1 Valid Classifi-cations	Cell 2 False Positives	Cell 3 False Positives	Cells 1-3
Borderline	Cell 4 False Negatives	Cell 5 Valid Classifi-cations	Cell 6 False Positives	Cells 4-6
Well-Adjusted	Cell 7 False Negatives	Cell 8 False Negatives	Cell 9 Valid Classifi-cations	Cells 7-9
Total in Class	Cells 1+4+7	Cells 2+5+8	Cells 3+6+9	Cells 1-9

Well-Adjusted sample mistakenly classified as being at risk. Cell 6 cases are also couples in the Well-Adjusted sample, but they have been misclassified as being borderline. A false positive rate can be calculated by considering the total number of cases in Cells 2, 3 and 6.

A false negative error is defined as when a test classifies a case as less "pathological" than it is in actuality. For our purposes, the false negative error occurs when the test classifies a married couple as being more adjusted than defined by the criterion. Cells 4, 7 and 8 represent false negative classifications. Cases in Cell 4 are married couples in the At Risk sample who have been incorrectly classified by the test as being borderline. Cell 7 cases are also couples in the At Risk sample, but they have been misclassified as being well-adjusted. Cell 8 cases are married couples in the Borderline sample mistakenly classified as being well-adjusted. A false negative rate can be calculated by considering the total number of cases in Cells 4, 7 and 8.

The classification matrices for Analysis II are listed in Table 10 in the same manner as they were for the first analysis.

From a univariate perspective, the results of Analysis II show that 18 of the predictor variables are significantly related to the criterion. A significant inverse relationship was found between the following measures and the criterion:

TABLE 10

CLASSIFICATION MATRICES FOR ANALYSIS II

MATRIX A: 5-VARIABLE COMPUTER MODEL

Classification by Function	Criterion Classification			Total N Classified by Function
	At Risk	Borderline	Well-Adjusted	
At Risk	29	7	1	37
Borderline	9	28	16	53
Well-Adjusted	5	10	70	85
Total in Class	43	45	87	175

Hit Rate = 72.6%
False Positive Rate = 18.2%
False Negative Rate = 27.3%

MATRIX B: 10-VARIABLE COMPUTER MODEL

Classification by Function	Criterion Classification			Total N Classified by Function
	At Risk	Borderline	Well-Adjusted	
At Risk	31	9	2	42
Borderline	8	27	15	50
Well-Adjusted	4	9	70	83
Total in Class	43	45	87	175

Hit Rate = 73.1%
False Positive Rate = 19.7%
False Negative Rate = 23.9%

TABLE 10 (CONTINUED)

MATRIX C: 15-VARIABLE COMPUTER MODEL

Classification by Function	Criterion Classification			Total N Classified by Function
	At Risk	Borderline	Well-Adjusted	
At Risk	32	10	1	43
Borderline	8	28	18	54
Well-Adjusted	3	7	68	78
Total in Class	43	45	87	175

Hit Rate = 73.1%
False Positive Rate = 22.0%
False Negative Rate = 20.5%

MATRIX D: 20-VARIABLE COMPUTER MODEL

Classification by Function	Criterion Classification			Total N Classified by Function
	At Risk	Borderline	Well-Adjusted	
At Risk	30	9	1	40
Borderline	8	29	16	53
Well-Adjusted	5	7	70	82
Total in Class	43	45	87	175

Hit Rate = 73.7%
False Positive Rate = 19.7%
False Negative Rate = 22.7%

TABLE 10 (CONTINUED)

MATRIX E: 28-VARIABLE COMPUTER MODEL

Classification by Function	Criterion Classification			Total N Classified by Function
	At Risk	Borderline	Well-Adjusted	
At Risk	30	8	1	39
Borderline	9	30	14	53
Well-Adjusted	4	7	72	83
Total in Class	43	45	87	175

Hit Rate = 75.4%
False Positive Rate = 17.4%
False Negative Rate = 22.7%

all Tharp scores except W II, all three DIPORE scores, and both Husband and Wife Stress Factors. Higher values on these measures are associated with correspondingly lower levels of adjustment.

A significant positive relationship was found between the following two scores and the criterion: R-C and Husband's age at marriage. Higher values on these measures are associated with correspondingly higher levels of adjustment. A significant curvilinear-type relationship was found between the following two measures and the criterion: Wife's age at marriage and Number of children. Borderline wives were found to be oldest at marriage while At Risk wives were youngest. Borderline marriages were found to have the most children while Well-Adjusted couples had the fewest.

Three of the 18 significant variables were demographic in nature. Of the 19 variables found significant in Analysis I all but two were found to be of significance in Analysis II.

From a multivariate perspective, only the first five variables selected by the step-wise analysis account for significant variance in the criterion groups. These five variables are H SX, Husband Stress Factor, Wife Stress Factor, R-C and W SL. Three of these scores were ranked in the top five in Analysis I. Combining these five variables into a multivariate classification model results in a respectable hit rate of 72.6 percent (see Table 10, Matrix A).

If the 10-variable classification model (Table 10, Matrix B) is used the hit rate is improved (73.1 percent) by only one case. However, the crucial false negative rate drops from 27.3 percent to 23.9 percent since the larger model misclassifies three fewer cases in a false negative direction.

The hit rate remains the same for the 15-variable model (Table 10, Matrix C) but there is improvement in the false negative rate (20.5 percent). There are slight gains in the overall hit rates for the 20-variable (Table 10, Matrix D) and 28-variable (Table 10, Matrix E) models. However, the false negative rate jumps up again to 22.7 percent for both these classification schemata. The false positive rate is best using the 28-variable model (17.4 percent) and worst using the 15-variable model (22.0 percent).

Though generally as accurate, the last two of the computer models are relatively inefficient when compared with the parsimony and cost savings of the three smaller models. While a case can be made for using the 15-variable model the insignificance of 10 F to Remove values would contraindicate such a decision. In sum, the 5- and 10-variable programs appear to be the most likely choices for a computer classification model. The former is most parsimonious while the latter elevates the hit rate in the process of decreasing the number of false negative errors.

Given the above findings, it was decided to cross-validate the classification program using both the 5-variable and 10-variable computer models.

Analysis II: Cross-Validation

The purpose of this analysis was to cross-validate the 5- and 10-variable classification models on a separate sample of At Risk, Borderline and Well-Adjusted couples. This was accomplished by applying the discriminant functions of these models to the appropriate data for the remaining At Risk (N=15), Borderline (N=15) and Well-Adjusted (N=30) cases.

Tables 11 and 12 show the classification matrices for these cross-validation runs. Inspection of the tables shows that the smaller model correctly classified one more case than the larger model. The hit rate for the former equals 56.7 percent while the latter hit on 55.0 percent of the total sample. These rates represent a shrinkage in overall hits of 15.9 percent and 18.1 percent, respectively, compared to the original standardizations.

The false negative rate is higher for the 10-variable model than for the 5-variable model on the cross-validation run. It is important to note which errors account for this rate differential. The 10-variable model generates 12 false negatives while the 5-variable model generates 10 false negatives. The extra cases in the former model are accounted for by two Borderline couples who have been mistakenly categorized as well-adjusted. This kind of false negative error

TABLE 11

CROSS-VALIDATION MATRIX OF ANALYSIS II 5-VARIABLE MODEL

Classification by Function	Criterion Classification			Total N Classified by Function
	At Risk	Borderline	Well-Adjusted	
At Risk	8	4	1	13
Borderline	5	8	11	24
Well-Adjusted	2	3	18	23
Total in Class	15	15	30	60

Hit Rate = 56.7%
False Positive Rate = 35.6%
False Negative Rate = 33.3%

TABLE 12

CROSS-VALIDATION MATRIX OF ANALYSIS II 10-VARIABLE MODEL

Classification by Function	Criterion Classification			Total N Classified by Function
	At Risk	Borderline	Well-Adjusted	
At Risk	8	4	1	13
Borderline	5	6	10	21
Well-Adjusted	2	5	19	26
Total in Class	15	15	30	60

Hit Rate = 55.0%
False Positive Rate = 33.3%
False Negative Rate = 40.0%

is the least severe of the three types of false negative misclassifications which can be made. That is, it would have been more "costly" if these couples had been drawn from the At Risk sample and had been classified by the test as borderline or well-adjusted. For this reason, the distinction between the two models' false negative rates is less than what might be initially inferred from looking only at the percentage differences.

The false positive rate is lower for the 10-variable model since this program classified one fewer Well-Adjusted couple as being borderline than did the 5-variable model.

Based on the standardization and cross-validation findings for Analysis II it can be concluded that both the 5-variable and 10-variable classification models are adequate devices for discriminating between At Risk, Borderline and Well-Adjusted criterion groups. Though these models classify at a level significantly above chance, their performance proficiency is not without considerable error.

Comparison of the Two Analyses

Results from the second analysis show that statistical techniques are not as ideal in discriminating between three kinds of marital criterion groups. Though the Analysis I hit rates are considerably higher than the Analysis II hit rates there is relative consistency in how the two analyses monitor the relationship between the predictor variables and the dependent measure. The predictors were ranked according

to their relative importance to the criterion in both programs--first from a univariate perspective (see Table 13) and then from a multivariate perspective (see Tables 4 & 8). Two Spearman rank-order correlation (r_s) coefficients were calculated to assess the degree of correspondence between these predictors over both analyses.

Highly significant correlations were found for both the univariate rank order of the predictors (r_s=.97; p<.001; df=26) and the multivariate rank order (R_s=.58, p<.002; df=26). These data show that the attenuation of hit rates in Analysis II is not due to a general breakdown in the relationships between the independent measures and the dependent variable. Rather, the classification task which considers the total adjustment continuum is considerably more complex, prone to error and demanding for a statistical model than the classification task based on only the extremes of that dimension.

TABLE 13

UNIVARIATE RANK ORDER IMPORTANCE OF PREDICTOR VARIABLES TO CRITERION

Predictor Variable	Analysis I Rank	Analysis II Rank
H II	7	8
H DR	8	9
H SL	4	3
H ER	9	10
H SX	2	1
W II	19	20
W DR	17	17
W SL	3	4
W ER	13	14
W SX	5	6
DAVE H	22	22
DAVE W	25	25
DAVE P	18	19
DIPORE H	12	12
DIPORE W	14	13
DIPORE P	11	11
R-C	10	7
Husband Stress Factor	1	2
Wife Stress Factor	6	5
Length of engagement	20	23
Husband's age at marriage	16	18
Wife's age at marriage	15	15
Length of marriage	27	27
Number of children	23	16
Number of hours husband employed	21	24
Total annual income	26	26
Husband's father's occupation	28	28
Wife's father's occupation	24	21

CHAPTER IV

THE CLINICAL APPROACH: METHOD

Rationale of Decision-Making Task

The second half of this investigation was addressed to an evaluation of a clinical decision-making task. The empirical and methodological foundation for this task was formulated using results from the previous section. The purpose was to appraise how well clinical judges could make predictions about married couples using the same three-group criterion as used by the statistical program. This phase of the research explored the hypotheses that decision-making tasks have teaching value and that clinicians are no more accurate than the actuary on a complex prediction task.

Findings already presented for Analysis II show no gross differences between the 5-variable and 10-variable classification models. Either model could have been used to generate the clinical decision-making task. However, a prediction task derived from the former model would present judges with information on only five measures, all of which are psychological in nature. The latter model permits construction of a judgment task which presents clinicians with information on 10 measures, including four which are demographic in nature.

The 5-variable judgment task would be more parsimonious than the 10-variable one. Nonetheless, it would be a fairly dull and unchallenging exercise since it presents the judges with only a limited pool of information from which to make their decisions. Since the presentation of demographic data was also important it seemed even more sensible to adapt the 10-variable classification model into the judgment task.

The Marriage Profile Summary

In order to simulate a clinical situation whereby judges could examine pertinent case material on married couples it was necessary to design the Marriage Profile Summary (MPS). The MPS presents the marriage-related data in a graphic form not unlike the Minnesota Multiphasic Personality Inventory (Dahlstrom & Welsh, 1960). It provides for a composite display of the 10 variables in a concise and easy-to-process fashion. The MPS form is shown in Appendix B.

Profiles for each of the 235 married couples were constructed. This was accomplished by transforming their scores on the six psychological variables--H DR, H SX, W SL, R-C, Husband Stress Factor and Wife Stress Factor--into standard scores and then into T scores. These data were plotted on the MPS grid. To the right of the grid information regarding the four demographic variables--Husband's age at marriage, Wife's age at marriage, Number of children and Wife's father's occupation--was also provided. The profiles were anonymous but did contain code numbers.

To facilitate the administration of the decision-making task the predictor data for the 175 standardization couples were printed on white MPS forms while blue MPS forms were used for the 60 cross-validation couples. For categorization purposes, gummed foil stars were used to identify the white MPS forms in terms of the three dependent criterion groups. Those white MPS forms representing Well-Adjusted couples were labeled with a green star. Similarly, the white forms for the Borderline couples were marked with a yellow star and the white forms for the At Risk couples were marked with a red star. Examples of these three types of adjustment profiles are contained in Appendix B. The blue MPS forms were purposely left unidentified because of the nature of the judgment task.

The Experimental Design

Judges

The sample of judges (Js) consisted of 36 individuals drawn from various areas in the University of Florida community. This sample represented three distinct levels of academic-professional training and experience.

The first set of 12 Js were enrolled as undergraduate students at the University and were involved with an introductory or more advanced psychology course at the time. Of these Js, seven indicated they were psychology majors while five indicated they were majoring in another discipline.

The mean number of psychology courses taken by these undergraduates was 5.75. Nine women and three men made up this sub-group within the total sample.

The next set of 12 Js were enrolled as graduate students at the University and were actively pursuing doctoral degrees in the clinical psychology training program. The mean number of years these Js had spent in graduate school was 2.33. Nine men and three women made up this sub-group within the total sample.

The final set of 12 Js were professionally engaged in clinical settings where mental health services were administered. All Js were full-time employees who were financially remunerated for their work. Ten of these professionals had earned the Ph.D., one had earned the Ed.D., and one was culminating his dissertation after a number of years in the field. Three Js worked as psychologists in the local Veteran's Administration hospital, three worked in the Psychology Clinic of the University's teaching hospital, three worked in the University Counseling Center, and three worked as therapists in the mental health division of the University Infirmary. The mean number of years'experience as reported by these Js was 5.58. Twelve men made up this sub-group within the total sample.

Basic Paradigm

The experimental design for the decision-making task is schematized in Table 14. Js from each of the three levels

TABLE 14

EXPERIMENTAL DESIGN SCHEMATIC

	Experience Level	Decision-Making Phase			
		I: Didactic Session		II: Judgment Session	
Treatment Conditions	Professional Clinicians (N=6)	1 2 3	4 5 6	1-A 2-A 3-A	4-B 5-B 6-B
	Graduate Students (N=6)	7 8 9	10 11 12	7-C 8-C 9-C	10-D 11-D 12-D
	Undergraduate Students (N=6)	13 14 15	16 17 18	13-E 14-E 15-E	16-F 17-F 18-F
Control Conditions	Professional Clinicians (N=6)			19-G 20-G 21-G	22-H 23-H 24-H
	Graduate Students (N=6)			25-I 26-I 27-I	28-J 29-J 30-J
	Undergraduate Students (N=6)			31-K 32-K 33-K	34-L 35-L 36-L

Each arabic numeral (1-36) represents one judge.
Each capital letter (A-L) represents the decision-making group which each judge was assigned to after making his individual decisions.

Group A: All Js employed at Veteran's Administration Hospital
Group B: All Js employed at University Counseling Center
Group G: All Js employed at University Hospital Psychology Clinic
Group H: All Js employed at Mental Health Division of University Infirmary

of experience were divided equally into one of two experimental conditions--treatment or control. These assignments were not made on a random basis due to problems of scheduling and previous time commitments. This was especially true for the professional Js who were involved in numerous clinical responsibilities. Since testing was conducted in teams of three, each of the four professional groups consisted of Js who were working at the same facility.

Treatment Conditions

Js selected as the treatment group were presented with a two-stage decision-making task spread over a two-day period. The first phase of the task was referred to as the Didactic Session while the second phase was referred to as the Judgment Session (see Table 14).

The Didactic Session involved more of a learning set than a performance orientation on J's behalf. He was presented with three separate folders containing the white MPS forms. Enclosed in one folder were the 43 At Risk profiles identified with red stars. The other two folders contained the 45 Borderline (yellow) and 87 Well-Adjusted (green) profiles, respectively. J's task was to study and analyze the profiles in each of the three adjustment groups so that he could develop an understanding of the relationships between the psychological-demographic data and the adjustment criterion.

The Judgment Session involved more of a performance orientation on J's behalf and was designed to appraise the validity of the relationships he had isolated during the Didactic Session. J was presented with a folder containing the 60 blue MPS forms (i.e., cross-validation sample). J's task was to determine the adjustment level for each of these new couples given the limited information on the profiles.

A more detailed description of these two sessions is given below.

Phase I: The Didactic Session

In this session J was presented with a 10-page booklet or manual entitled "Marital Adjustment in a College Population: A Learning Exercise" (see Appendix C). The cover page required J to provide some biographical information. Also contained on this face sheet was an innocuous code number used to identify the treatment and control groups.

Page two delineated the general instructions and explained the purpose and basic outline of the study. The next section, pages three through six, provided more specific information about the learning exercise. This information consisted of:

1. A general definition of marital adjustment and a conceptual demarcation of the adjustment dimension into the three criterion categories--At Risk, Borderline and Well-Adjusted.

2. An explanation of the 10 psychological and biographical predictors available for each couple and displayed on the profile material.

3. A paragraph instructing J on how to read the MPS grid and the appropriate procedure for interpreting negative and positive scores. J was also informed of the standardization procedures by which the profile plots were determined. Mean scores for the four biographical predictors were also provided since these variables were not presented on the T score scale.

4. A statement indicating that previous statistical analyses had shown the 10 variables to be highly accurate, but not perfect, in predicting the adjustment levels of those couples whose profiles J was going to review. Therefore, J was informed that linear combinations of data were not always successful at classifying adjustment.

J was also asked to estimate his pre-treatment level of expertise concerning the psychological and sociological dynamics of college marriage, using a rating scale of 1-100.

Page seven contained a few instructional paragraphs which prepared J for the actual task. The three packets of MPS material were explained and J was told to scrutinize as many as or as few of the data he cared to look over. While studying the profiles J was instructed to map out in his cognitive system any relationships which seemed to hold

between the predictors and the criterion. After this analysis J was asked to turn to the remaining section of the manual where various forms of feedback were required.

Pages eight through 10 contained a series of evaluation measures which were used to monitor the effects of the Didactic Session and the kinds of interpretations which J had made from the data. The four kinds of feedback were as follows:

1. J provided a list of decision rules which he felt could be used to accurately classify a majority of the profiles into the three adjustment groups. These could be in the form of rules of thumb, cookbook formulas or if-then statements.
2. J ranked the 10 MPS variables in terms of their discriminable order of importance to marital adjustment.
3. J estimated the importance of the 10 variables to the criterion by assigning each a weight from 0-100. J was instructed to assign the weights so they represented the relative degree of importance of one variable to another. It was also necessary that the sum of the assigned weights total 100.
4. J estimated the base rates for At Risk, Borderline and Well-Adjusted marriages in the college population, assuming that the white MPS data were representative of such a population.

Space was also provided for any comments J cared to make about the Didactic Session. Finally, J was instructed as to the time and place of the next day's Judgment Session.

Phase II: The Judgment Session

In this session J was presented with a six-page booklet (see Appendix C). Besides the cover page this manual consisted of a set of instructions for the individual decision-making task. J was informed that the 60 blue MPS forms were derived from the same standardization group as the previous day's sample. His task was to classify each profile into one of the three marital criterion groups. J was advised to make use of any of the information which he had provided in Phase I as well as any new insights he had gained. The booklet from the Didactic Session was returned to J for this purpose.

For each case classified J was told to rate his confidence in that decision on a scale from 50-100 percent. More certain decisions were to be assigned higher percentages. If J were positive of his decision he marked 100 percent. If he were only guessing he marked 50 percent. Other confidence scores were assigned in intervals of 10.

Pages three through five contained a tabulation chart on which J scored his decisions. On the final page of the booklet space was provided for any comments J cared to make.

Control Conditions

Js selected as the control group were presented with a single-stage decision-making task. In actuality, the control conditions were analogous to the Judgment Session in Phase II of the treatment procedures (see Table 14). That is, control Js were presented only with the 60 blue MPS forms and asked to classify these profiles into the adjustment groups without ever having the advantage of viewing the 175 white MPS forms. Stated in other words, treatment Js were exposed to both standardization and cross-validation profiles while control Js were exposed to just the cross-validation group. In this manner, the control conditions allowed for an assessment of the teaching effects produced by the Didactic Session in Phase I of the treatment paradigm.

J was presented with a 13-page booklet (see Appendix C) with the same previously mentioned title. This version of the manual was a composite of the two booklets presented to the treatment group. The cover sheet was followed by a page of general instructions. The next sections, pages three through six, provided descriptions of the criterion and the predictor variables. Except for changes in some orientation passages this material was an exact duplication of the information provided the treatment group. Similarly, J estimated his pre-task level of expertise in the subject area.

Page seven contained a few instructional paragraphs which prepared J for the actual task. J was told to

scrutinize as many or as few of the data as he cared to on the blue MPS forms. After studying the profiles J was asked to provide feedback about the predictor-criterion relationships based on information from the previous pages and his clinical insights, hypotheses and hunches about the variables. Pages eight and nine contained the same kinds of evaluation measures described above for the treatment conditions--i.e., decision rules, rank order ratings, assigned weights, and base rate percentages.

The second and major part of the judgment task was also described on page eight. J was instructed to classify the 60 blue profiles into one of the three adjustment levels and to assign confidence scores to his decisions. The procedures adopted here were the same as those used by the treatment group in making its classifications. Tabulation charts were provided on pages 10 through 12. On the final page of the booklet space was provided for any pertinent comments.

Group Decision Making

After completion of the individual judgment task J combined together with the other two judges present in the room to make group decisions about the same data (see Table 14). J was instructed to gather together with the other persons on his team to share his insights and cognitive rules in a roundtable fashion. The group was given time to discuss the relationships between the predictors and the criterion and

was then asked to reclassify the profiles based on group agreement. Group confidence ratings were also assigned as determined by the decision-making team. Separate tabulation charts were provided for this group task.

After this second decision-making exercise J was asked to estimate his post-task level of expertise concerning the psychological and sociological dynamics of college marriage. Once again, a 100-point scale was used.

Independent Variables

A review of Table 14 shows that a 3x2 experimental design was developed to assess the role of two independent variables--experience level and experimental conditions--affecting J's performance in predicting marital adjustment.

The first variable was divided into three different levels--professional, graduate and undergraduate. The second variable was divided into two different levels--the treatment group, exposed to both didactic (standardization) and judgment (cross-validation) sessions and the control group, exposed to only the judgment session.

A third independent variable of importance to the total design, whose role was evaluated in separate analyses, was individual versus group decision making. As used in the present clinical judgment task this variable was divided into two dichotomous levels--individual decisions and group decisions.

Basic Dependent Measures

Five dependent measures were used to assess J's performance on the decision-making task. These measures were derived from J's responses to the 60 blue profiles in the Judgment Session. Most of these measures have been mentioned in previous sections and will be summarized briefly below.

Three measures were isolated to evaluate J's accuracy in the decision-making task. The total hit rate was determined by computing the proportion of cases correctly classified by J. The false positive rate was determined by computing the proportion of cases whose adjustment level was underestimated by J. The false negative rate was determined by computing the proportion of cases whose adjustment level was overestimated by J.

One measure was isolated to evaluate J's confidence in his own decision-making skills. The mean confidence score was determined by computing the average percentage rating assigned by J to the 60 cases.

A final measure was isolated to evaluate J's appropriateness in the decision-making situation. Decision-making behavior is appropriate when a person is either accurate and assigns high confidence ratings to his judgments or is inaccurate but assigns low confidence ratings to his judgments. The appropriateness score was determined by computing a Pearson product-moment correlation coefficient between the accuracy of J's 60 decisions and the confidence scores he

assigned to them. When J was correct in his prediction he was given an accuracy value of three (3) for that decision. When J's prediction was off by one adjustment level he was given an accuracy value of two (2). When J's prediction was off by two adjustment levels he was given an accuracy value of one (1).

These measures were applied to both the individual and group data.

Additional Dependent Measures

J's responses to other sections in the test booklets provided ancillary kinds of information worth considering. These five additional measures have been mentioned earlier in this chapter. Their relevance to J's performance will be outlined below.

The decision rule data supplied by J allow for a content analysis of the kinds of cognitive processes he used to define the relationships between the predictors and the criterion.

The rank order data provide a gross measure of how well J perceived the empirical relationships as compared with the multivariate analysis. A correlation coefficient can be computed to measure the degree of correspondence between J's rank order and the computer model's rank order. Higher correlations would indicate that J had learned something about the relationships between predictors and criterion.

The weight data provide a more molecular measure of how much J learned from his analyses of the profiles. These weights can be compared with the program's weights to note how J's multivariate view of the 10 factors compared with the actual statistical analyses of the data.

The base rate data provide a measure of how good Js in the treatment group were at simple counting procedures. To have accurately estimated the population base rates, J needed only to determine the proportion in each criterion group to the total sample. This information would have been extremely helpful to him in the Judgment Session of the decision-making task.

Finally, the pre- vs. post-test level of expertise data provide a self-evaluative measure of J's subjective experience of the decision-making task and the amount of learning he believes took place.

Summary of Procedure

Js were tested in small groups of three. Treatment Js were tested over two sessions, each lasting approximately 90 mins. In the Didactic Session J was instructed to work on the MPS data and test booklet without consulting any of the other judges in the room. Upon leaving the Phase I session J was also asked not to discuss his participation in the project with anyone else. In the next day's Judgment Session J first worked on the MPS data and test booklet individually. At conclusion, he united with the other judges

for the group decision-making task. Control Js were tested in one session lasting approximately two hours. Similarly, J proceeded to first work on the task as an individual and then as a group member. In total, 12 groups were tested in 18 sessions.

CHAPTER V

THE CLINICAL APPROACH: RESULTS

Basic Research Questions

The primary statistical analyses were addressed to four basic research questions discussed in Chapters I-IV. These issues can be briefly categorized as follows:

1. Importance of Didactic Training. Will judges who have been presented with a didactic experience designed to teach them material about marital adjustment fare better in making predictions about couples than judges who are not afforded a similar learning opportunity? Will trained judges differ in confidence and appropriateness from their non-trained counterparts?

2. Role of Experience. Will judges who have had experience making real-life clinical decisions be more accurate in appraising marital adjustment than inexperienced, naive judges? Will they differ in their confidence, appropriateness and the nature of the errors which they make? If any differences exist will they be influenced by common training procedures?

3. Individual versus Group. Do human judges become more accurate when they unite in a team and pool their individual resources to make group decisions about marital adjustment?

4. Man versus Machine. How do human judges compare with a computerized, statistical methodology in making predictions about the marital adjustment of college marrieds? Can humans surpass the performance of the machine? Do they make similar kinds of errors with less frequency? The man versus machine problem is conditionally related to the other issues and, in essence, is the global question behind this research.

The results will be described with regard to these four research issues. Of course, there is obvious overlap in how the findings can be applied to these questions, and the breakdown presented below is organized for the purpose of conceptual clarity.

Hypotheses and Major Statistical Analyses

Each of the basic research questions can be phrased more operationally in terms of a set of hypotheses. Table 15 lists the four sets of hypotheses which were tested using various statistical procedures. Five Two-Way Analyses of Variance (ANOVAs) were generated to determine the main treatment effects of the experimental conditions and experience level variables. These ANOVAs are summarized in Tables 16-20 and have relevance to the first 10 hypotheses. The means

TABLE 15

HYPOTHESES USED TO EVALUATE BASIC RESEARCH QUESTIONS

Issue		Hypothesis
Importance of Didactic Training	1.	Hit rates on the judgment task will be significantly higher for treatment Js than for control Js. This relationship should hold over all three experience levels.
	2.	False positive rates on the judgment task will be significantly lower for treatment Js than for control Js. This relationship should hold over all three experience levels.
	3.	False negative rates on the judgment task will be significantly lower for treatment Js than for control Js. This relationship should hold over all three experience levels.
	4.	Confidence ratings on the judgment task will be significantly higher for treatment Js than for control Js. This relationship should hold over all three experience levels.
	5.	Appropriateness scores on the judgment task will be significantly higher for treatment Js than for control Js. This relationship should hold over all three experience levels.
Role of Experience	6.	Level of experience will have no significant effect on the hit rates of either treatment or control Js on the judgment task.
	7.	Level of experience will have no significant effect on the false positive rates of either treatment or control Js on the judgment task.
	8.	Level of experience will have no significant effect on the false negative rates of either treatment or control Js on the judgment task.

TABLE 15 (CONTINUED)

Issue		Hypothesis
Role of Experience (cont.)	9.	Level of experience will have no significant effect on the confidence ratings of treatment Js on the judgment task. However, more experienced control Js should report significantly lower confidence ratings.
	10.	Level of experience will have no significant effect on the appropriateness scores of treatment Js on the judgment task. However, more experienced control Js should obtain significantly higher appropriateness scores.
Individual versus Group	11.	When functioning as a group, level of experience will have no significant effects on the accuracy, confidence, and appropriateness measures of treatment Js.
	12.	When functioning as a group, level of experience will have no significant effects on the accuracy measures of control Js. However, when functioning as a group, more experienced control Js should report significantly lower confidence ratings and significantly higher appropriateness scores.
	13.	When functioning as a group treatment Js will be significantly higher on the accuracy (lower on false positive and false negative rates), confidence and appropriateness measures than when functioning as individuals. These relationships should hold over all three experience levels.
Man versus Machine	14.	The hit rate for treatment Js on the judgment task will not differ significantly from the computer model's cross-validation hit rate of 55 percent. This relationship should hold over all three experience levels.

TABLE 15 (CONTINUED)

Issue		Hypothesis
	15.	The false positive rate for treatment Js on the judgment task will not differ significantly from the computer model's cross-validation false positive rate of 33.3 percent. This relationship should hold over all three experience levels.
Man versus Machine (cont.)	16.	The false negative rate for treatment Js on the judgment task will not differ significantly from the computer model's cross-validation false negative rate of 40.0 percent. This relationship should hold over all three experience levels.
	17.	When functioning as a group, the hit rate for treatment Js on the judgment task will be significantly higher than the computer model's hit rate. Similarly, the false positive and false negative rates for treatment Js will be significantly lower than the analogous rates for the computer model. These relationships should hold over all three experience levels.

TABLE 16

ANALYSIS OF VARIANCE OF THE MAIN TREATMENT EFFECTS AS MEASURED BY TOTAL HIT RATE (THR)

Source	df	MS	F
Experience Level (Rows)	2	0.000	0.19
Experimental Conditions (Columns)	1	0.003	1.21
Interaction (Row x Column)	2	0.001	0.31
Error (Within Cells)	30	0.003	
Total	35		

TABLE 17

ANALYSIS OF VARIANCE OF THE MAIN TREATMENT EFFECTS AS MEASURED BY FALSE POSITIVE RATE (FPR)

Source	df	MS	F
Experience Level (Rows)	2	0.005	0.41
Experimental Conditions (Columns)	1	0.000	0.00
Interaction (Row x Column)	2	0.022	1.84
Error (Within Cells)	30	0.012	
Total	35		

TABLE 18

ANALYSIS OF VARIANCE OF THE MAIN TREATMENT EFFECTS AS MEASURED BY FALSE NEGATIVE RATE (FNR)

Source	df	MS	F
Experience Level (Rows)	2	0.008	0.60
Experimental Conditions (Columns)	1	0.018	1.41
Interaction (Row x Column)	2	0.026	2.06
Error (Within Cells)	30	0.013	
Total	35		

TABLE 19

ANALYSIS OF VARIANCE OF THE MAIN TREATMENT EFFECTS AS MEASURED BY MEAN CONFIDENCE SCORE (MCS)

Source	df	MS	F
Experience Level (Rows)	2	0.051	7.04*
Experimental Conditions (Columns)	1	0.012	1.64
Interaction (Row x Column)	2	0.005	0.72
Error (Within Cells)	30	0.007	
Total	35		

*$p < .01$

TABLE 20

ANALYSIS OF VARIANCE OF THE MAIN TREATMENT EFFECTS AS MEASURED BY APPROPRIATENESS SCORE (AS)

Source	df	MS	F
Experience Level (Rows)	2	0.010	0.69
Experimental Conditions (Columns)	1	0.036	2.51
Interaction (Row x Column)	2	0.008	0.57
Error (Within Cells)	30	0.014	
Total	35		

and standard deviations for each of the primary dependent variables are compared across the five treatment levels--professional clinicians, graduate students, undergraduate students, treatment judges and control judges--in Table 21. Another series of ANOVAs (Table 22) was applied to the group decision-making data in order to assess Hypotheses 11-13. Simple t-tests as well as information from Table 21 and a second table of means for the group decision-making task (Table 23) were used in evaluating the final four hypotheses.

Importance of Didactic Training; Role of Experience

Table 16 presents the ANOVA of the main experimental effects as measured by the Total Hit Rate (THR). Experimental conditions did not significantly influence performance on this dependent variable although, as expected, treatment Js had a slightly greater proportion of correct judgments (54.2 percent) than control Js (52.3 percent). This result appears to disconfirm Hypothesis 1. Experience level also did not significantly influence the THR performance. Differences between the three experience groups were minimal under both treatment and control conditions. This result confirms Hypothesis 6 which had suggested no relationship between experience and prediction accuracy. Also, no interaction effect on the dependent variable was found.

TABLE 21

COMPARISON OF THE DIFFERENT TREATMENT GROUPS ON THE PRIMARY DEPENDENT VARIABLES

Treatment Group	N	Measure									
		Total Hit Rate		False Positive Rate		False Negative Rate		Mean Confidence Score		Appropriateness Score	
		$\bar{M}$	SD	$\bar{M}$	SD	$\bar{M}$	SD	$\bar{M}$	SD	$\bar{M}$	SD
Professional	12	53.2	6.1	36.8	12.2	39.7	11.2	72.4	8.5	.180	.136
Graduate	12	53.9	4.9	36.3	12.4	37.8	13.9	75.7	7.0	.236	.112
Undergrad.	12	52.6	3.6	40.0	7.7	34.7	9.4	85.0	9.9	.217	.111
Treatment	18	54.2	5.3	37.6	10.0	35.2	9.4	79.5	9.6	.242	.101
Control	18	52.3	4.3	37.8	11.8	39.6	13.2	75.9	10.1	.179	.131
All Judges	36	53.2	4.9	37.7	10.8	37.4	11.5	77.7	9.9	.211	.119
Computer Model		55.0		33.3		40.0					
Js > Computer		10		12		24					
Treatment		7		5		13					
Control		3		7		11					

The ANOVA of the main experimental effects as measured by the False Positive Rate (FPR) is outlined in Table 17. There were no significant row, column or interaction effects. Since Js in the treatment group did not have lower FPRs than their counterparts in the control group, Hypothesis 2, which had predicted such a difference, is disconfirmed. The results confirm Hypothesis 7 which had suggested no relationship between experience and false positive misclassifications.

Table 18 outlines the ANOVA of the main experimental effects as measured by the False Negative Rate (FNR). Here too, there were no significant row, column or interaction effects. Although these results largely disconfirm Hypothesis 3, treatment Js were generally better predictors as measured by this statistic than were control Js (35.2 percent versus 39.6 percent). This discrepancy was most apparent at the graduate level of experience (30.6 percent versus 45.0 percent). Surprisingly, this relationship reversed itself at the undergraduate level (36.7 percent versus 32.8 percent). Although there appears to be a minimal trend toward some sort of interaction phenomenon, the lack of any overall statistically significant effects produced by the experience variable confirms Hypothesis 8. This means that on the average, the number of false negative misclassifications tended to be the same over all three experience levels.

The ANOVA of the main experimental effects as measured by the Mean Confidence Score (MCS) is outlined in Table 19. Experience level was found to have a significant effect on MCS while no experimental conditions or interaction effects were obtained. The results are presented graphically in Figure 1.

Inspection of the graph shows an inverse relationship between experience level and confidence. This relationship was most pronounced under the control conditions but less so under treatment conditions, since in the latter, confidence ratings by the professional and graduate groups were virtually the same.

To analyze the data more thoroughly the Newman-Keuls test (Kirk, 1968) was applied to evaluate the differences between group means. The results of these comparisons are given in Appendix D. The undergraduates were significantly more confident in their decisions than either the graduates or professional clinicians. However, these differences fall short of statistical significance when comparing only treatment groups and in the undergraduate/graduate comparison of control Js.

These findings suggest that inexperience led to greater confidence in the decision-making situation. However, using didactic procedures, the relative differences between inexperienced judges and their more experienced counterparts became less noticeable. This shrinkage is due in part to

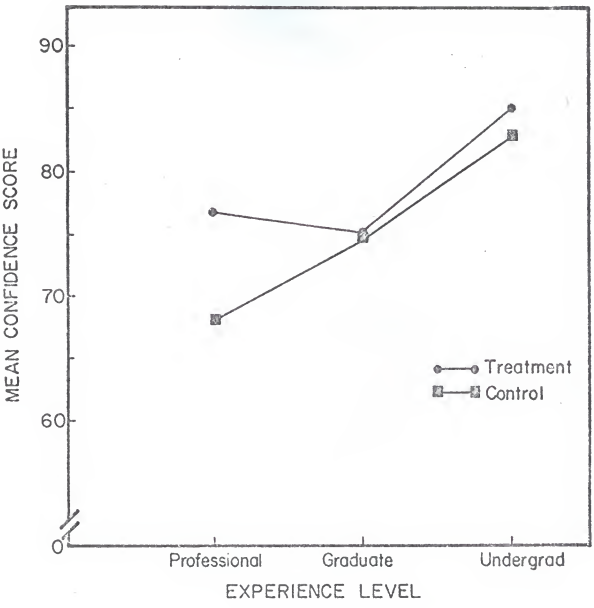


Figure 1. Judges' Mean Confidence Score (MCS) as a Function of Level of Experience

professional Js in the treatment group being more confident in their prediction than professional Js in the control group, even though this difference was not statistically significant. The didactic procedures did not have a similar elevating effect on either of the two student groups.

These results disconfirm Hypothesis 4 suggesting that, over all Js, there is no relationship between confidence and the type of experimental conditions under which decisions were made. However, for the most part, the data confirm Hypothesis 9 which had predicted differential experience effects in the control group but not in the treatment group.

Table 20 outlines the ANOVA of the main experimental effects as measured by the Appropriateness Score (AS). No significant row, column or interaction effects were found. However, there was a trend toward significance in the experimental conditions variable and the results are graphically presented in Figure 2. The treatment group, on the whole, tended to be more appropriate than the control group and this difference was most prominent for the professional Js. The mean appropriateness score was highest for the graduate students in the treatment condition. These findings suggest a trend toward confirmation of Hypothesis 5. The first half of Hypothesis 10 is confirmed but the second half

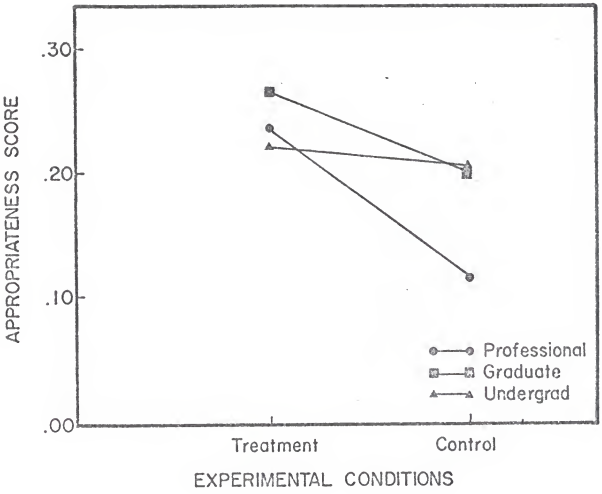


Figure 2. Judges' Appropriateness Score (AS) as a Function of Experimental Conditions of Decision-Making Task

is disconfirmed, which means that there was no relationship between experience and appropriateness in either of the two experimental conditions.

Individual versus Group

In order to test the next set of hypotheses it was first necessary to conduct similar analyses of variance on the group decision-making data. These ANOVAs are summarized in Table 22. The mean scores for the different treatment levels are listed in Table 23.

As can be seen, only one treatment effect was found across all five dependent measures--a finding which matches the results of the individual analyses. For the groups, too, experience level was found to have a significant ($p<.025$) influence on the MCS with the undergraduate groups most confident and the professional groups least confident. A Newman-Keuls analysis (see Appendix D) was conducted to further evaluate MCS differences. Overall, both the professional and graduate groups were significantly different from the undergraduate groups in their confidence ratings, but not significantly different from each other. However, these relationships were statistically more pronounced under treatment conditions than under control conditions.

These findings generally confirm Hypothesis 11 which was addressed to the role of experience in group decision

TABLE 22

ANALYSES OF VARIANCE OF THE MAIN TREATMENT EFFECTS ON THE GROUP DECISION-MAKING DATA

Source	df	Total Hit Rate (THR) MS	F	False Positive Rate (FPR) MS	F	False Negative Rate (FNR) MS	F	Mean Confidence Score (MCS) MS	F	Appropriateness Score (AS) MS	F
Experience Level (Rows)	2	.001199	0.82	.000660	0.08	.001106	0.28	.018420	9.95**	.000156	0.02
Experimental Conditions (Columns)	1	.000200	0.14	.005002	0.60	.005985	1.54	.002977	1.61	.000176	0.03
Interaction (Row x Column)	2	.000282	0.19	.003154	0.38	.005919	1.52	.001100	0.59	.000047	0.01
Error (Within Cells)	6	.001456		.008372		.003894		.001851		.006523	
Total	11										

**p < .025

TABLE 23

MEAN SCORES FOR THE GROUP DECISION-MAKING TASK ON THE PRIMARY DEPENDENT VARIABLES

Treatment Level	N	Measure				
		Total Hit Rate (THR)	False Positive Rate (FPR)	False Negative Rate (FNR)	Mean Confidence Score (MCS)	Appropri-ateness Score (AS)
Professional	4	52.9	36.1	40.0	72.7	.281
Graduate	4	56.3	33.9	36.7	75.0	.279
Undergrad.	4	53.8	36.1	38.3	85.4	.291
Treatment	6	53.9	37.4	36.1	79.3	.287
Control	6	54.7	33.3	40.6	76.1	.280
All Groups	12	54.3	35.4	38.3	77.7	.284
Computer Model		55.0	33.3	40.0		
Groups > Computer		3	4	5		
Treatment		1	1	3		
Control		2	3	2		

making for those judges exposed to the didactic procedures. Except for the confidence variable there was no relationship between experience and the five dependent measures.

Hypothesis 12, which was addressed to the role of experience in group decision making for those judges not exposed to the didactic procedures, is only partially confirmed. While there is a trend to suggest that more experienced control groups reported a lower MCS, no such trend exists with regard to the appropriateness measure.

Hypothesis 13 was evaluated by comparisons between the Individual and Group ANOVAs, made by pooling their respective MS_{error} values to obtain a best estimate of population variance. F tests for homogeneity of ANOVA variances were not significant and are summarized in Table 24. Each of the 3x2 cells for the Individual ANOVAs was compared with its Group ANOVA counterpart to evaluate for mean differences. These matches generated 30 simple t-tests across the five dependent measures of interest and the results are summarized in Table 25. No significant differences were found in any of these comparisons and the only suggested trend was that professional judges under control conditions may have been more appropriate making judgments as a group rather than as individuals. These results disconfirm Hypothesis 13.

TABLE 24

COMPARISON OF INDIVIDUAL AND GROUP ANOVAS FOR HOMOGENEITY OF VARIANCE

Dependent Measure	Individual ANOVA		Group ANOVA				
	MS	df	MS	df	F	S_p^2	est. $\sigma_{diff.}$
Total Hit Rate	.0030	30	.0015	6	2.00	.0028	.0432
False Positive Rate	.0120	30	.0084	6	1.43	.0114	.0872
False Negative Rate	.0130	30	.0039	6	3.33	.0115	.0876
Mean Confidence Score	.0070	30	.0019	6	3.68	.0062	.0643
Appropriateness Score	.0140	30	.0065	6	2.15	.0128	.0924

formula for pooled variance:

$$S_p^2 = \frac{\gamma_i MSE_i + \gamma_g MSE_g}{\gamma_i + \gamma_g}$$

formula for estimated standard error of the difference:

$$\text{est. } \sigma_{diff.} = \sqrt{S_p^2\left(\frac{1}{N_i} + \frac{1}{N_g}\right)}$$

N_i = number of observations in each Individual ANOVA cell
= 6 (note: $\gamma_i = N_i - 1$)
N_g = number of observations in each Group ANOVA cell
= 2 (note: $\gamma_g = N_g - 1$)

TABLE 25

CELL MEAN SCORES AND t TEST DIFFERENCES FOR THE INDIVIDUAL AND GROUP DECISION-MAKING TASKS

Cell Characteristics	Dependent Measure	Individual Mean	Group Mean	t
Professional-Treatment	THR	54.7	53.4	.30
	FPR	34.8	35.6	-.09
	FNR	38.3	40.0	-.19
	MCS	76.6	73.7	.45
	AS	.238	.289	-.55
Professional-Control	THR	51.7	52.6	-.21
	FPR	38.9	36.7	.25
	FNR	41.1	40.0	.13
	MCS	68.3	71.8	-.54
	AS	.121	.273	-1.63
Graduate-Treatment	THR	53.9	55.9	-.46
	FPR	41.1	38.9	.25
	FNR	30.6	30.0	.07
	MCS	75.8	75.3	.08
	AS	.266	.280	-.15
Graduate-Control	THR	53.9	56.7	-.65
	FPR	31.5	28.8	.31
	FNR	45.0	43.4	.18
	MCS	75.5	74.7	.12
	AS	.207	.278	-.77
Undergraduate-Treatment	THR	53.9	52.6	.30
	FPR	37.1	37.8	-.08
	FNR	36.7	38.3	-.18
	MCS	86.1	88.9	-.44
	AS	.223	.293	-.76
Undergraduate-Control	THR	51.4	55.0	-.83
	FPR	43.0	34.5	.97
	FNR	32.8	38.4	-.64
	MCS	83.9	82.0	.30
	AS	.210	.288	-.84

Man versus Machine

The THR grand mean for all 36 individual Js was 53.2 percent (see Table 21) as compared with 55.0 percent for the computer model. Since there were no row or column effects influencing this measure, all Js were collapsed into one group in order to make a statistical comparison with the computer model. This was done by a simple two-tailed t-test where the machine's hit rate was considered to be the expected value at which human judges would perform. In such a fashion the null hypothesis, H_o: THR = 55.0, was tested against the alternative hypothesis, H_1: THR $\neq$ 55.0. Using this procedure the obtained THR grand mean was found to be significant at the $p<.05$ level ($t=2.17$, $df=35$). All control Js were selected out and a similar comparison was made between only the treatment Js and the computer model. No significant statistical difference was found ($t=0.67$, $df=17$). This result generally confirms Hypothesis 14. Of the 10 judges who beat the actuary, seven came from the treatment group. Another three treatment Js had the same number of correct classifications as did the computer.

The FPR grand mean for all individual Js was 37.7 percent as compared with 33.3 percent for the computer model. Similar t-test procedures as described above were adopted and this difference was found to be significant at

the $p<.02$ level ($t=2.46$, $df=35$). However, comparison of the treatment group mean (37.6) and the computer's FPR showed that this difference only approached significance ($t=1.87$, $df=17$; $p<.10$). These results would appear to partially disconfirm Hypothesis 15. While treatment Js on the professional level did not depart noticeably from the actuarial rate ($\bar{M}=34.8$) their graduate student ($\bar{M}=41.1$) and undergraduate ($\bar{M}=37.1$) counterparts were obviously more prone to underestimating the level of marital adjustment on the cross-validation sample. Although 12 Js had fewer false positive errors than did the actuary, only five came from the treatment group. Another three Js, all from the treatment group, tied the computer on this statistic.

The FNR grand mean for all individual Js was 37.4 percent as compared with 40.0 percent for the computer model. The t-test comparison between judges and actuary was not significant ($t=1.36$, $df=35$). However, the performance of the treatment judges ($\bar{M}=35.2$) was statistically different from the computer's accuracy ($t=-2.18$, $df=17$; $p<.05$).

Of the 24 judges who beat the actuary on this statistic, 13 came from the treatment group. Another two Js, one from the treatment group, had the same number of false negative errors as did the computer. These results provide no support for Hypothesis 16. Instead, the trend is in the reverse direction, which would suggest that clinical judges in the

treatment group surpassed (rather than fell short of) the computer model's prediction performance as measured by the FNR statistic.

The THR grand mean for all 12 groups (54.3 percent; see Table 23) was not significantly different from the computer's hit rate (t=-.71, df=11). Although the treatment groups were even less accurate ($\bar{M}$=53.9) this difference was also insignificant (t=-1.08, df=5). All told, only three groups failed to match the computer THR performance; two of these were treatment groups.

The FPR grand mean for all groups (35.4) was not significantly different from the similar computer error rate (t=.99, df=11). The treatment group FPR (37.4) was sufficiently different to suggest a trend toward statistical significance (t=2.07, df=5; p<.10). This finding indicates that treatment judges, acting as a group, were more prone to misclassifying the married couples in a negative direction than was the computer. This difference tended to be most pronounced for the graduates and least pronounced for the professionals. All told, half the groups failed to match the computer FPR performance; four of these were treatment groups.

The FNR grand mean for all groups (38.3) was not significantly different from the similar computer error rate (t=-.91, df=11). The treatment groups' FNR (36.1) approached significance using a one-tailed test (t=1.66,

df=5; p<.10). This finding indicates that treatment judges acting as a group were less prone to misclassifying the married couples in a positive direction than was the computer. As with the FPR, this difference tended to be most pronounced for the graduates and least pronounced for the professionals. All told, only four groups failed to match the computer FNR performance; only one of these was a treatment group.

These comparisons between group decision making and the machine generally disconfirm Hypothesis 17 which had predicted greater proficiency for the human judges across the board. However, the assertion that the group false negative rate is less than the analogous computer error was not disconfirmed by the data.

Other Findings

While the above results are most crucial to this study there are a series of supplementary findings which may also be of interest and are reported in this section.

Content Analysis of Decision Rules

The different kinds of decision rules which Js provided were evaluated by the author using a simple content analysis procedure. Each rule was classified into one of six categories depending on the degree of specificity or amorphousness which characterized the statement. These categories were coded as follows:

1. Quantitative rules. These are specific mathematical statements which are written in terms of some sort of numerical equation. They are cookbook-type decision rules which explicitly define how the predictor cues can be used to assign couples to one of the three criterion groups. These rules can be stated in terms of the specific predictors themselves or, more generally, in terms of a group or number of predictors. The nature of these rules is such that they can be applied to the data in an objective, systematic fashion.

2-3. Qualitative rules. These are less definitive statements which posit some type of relationship between predictors and the criterion but suggest no algebraic functions by which classifications can be made. These rules are more generalized and probabilistic. They are usually phrased in the language of probabilities, likelihoods, correlations, positive and negative relationships, etc. The nature of these rules is such that they can only be applied to the data in a subjective, relative fashion since they are dependent on the individual judge's variance in interpretation. In this sense, the judge is not providing a rule but is suggesting a hedge which allows him enough degrees of freedom to vary in his judgments.

There are two sub-types within this category. Clear qualitative rules (2) are easily understood and suggest a definite relationship between predictors and criterion.

Vague qualitative rules (3) are less rigid and often cannot be applied to the data in a regular fashion. They may be stated in an unclear, ambivalent manner such that classifications can be made into more than one of the criterion groups, depending on the relative merits of the profile. Or they may be stated in an even more vague, ambiguous fashion --so idiosyncratic and loosely phrased that they can only be used privately by the judge who presents them.

4. Guidelines. These are statements which suggest cues, warning signs or indicators which the judge uses to focus on the pertinent aspects of the data. Usually, they suggest comparisons of data which are necessary or material which should be included or excluded from consideration in making a decision.

5. Comments. These are statements made by the judge which reflect some of his thinking about the data. They cannot be used to make classifications but suggest some attempt by the judge to communicate his recognition of different empirical relationships in the data.

6. Requests. These are statements by which the judge asks for more information than that provided on the profile. In this sense, the judge seems to be indicating that he is having difficulty using the available data to generate appropritate decision rules.

Appendix E contains the decision rules provided by each of the 36 Js in the study and shows how each was coded.

Table 26 presents the number and types of rules generated over treatment levels and within treatment groups. As was to be expected, treatment Js provided a substantially larger number of rules than did control Js. The former group listed quantitative rules most frequently (49.4 percent) while the latter group provided clear qualitative rules with the greatest frequency (42.2 percent). Control Js were also more than three times as likely to give comments or requests than were treatment Js (23.4 percent as compared with 7.1 percent).

There were no striking differences between the total number of rules generated by the three experience groups. The greatest number of quantitative rules were provided by graduate students while the fewest were generated by undergraduates. However, when comparing only treatment Js, there was no noticeable difference between graduates and professionals. The professional clinicians had the greatest tendency to suggest vague qualitative rules. The same group also made the greatest number of comments or requests. Graduate students, meanwhile, provided the fewest number of comments or requests but stated the most guidelines.

In sum, the treatment conditions of the design seemed to increase the likelihood of rule-generating by judges. These same conditions also increased the probability that rules would be specific, numerical statements rather than generalized, qualitative statements. Of the three levels

TABLE 26

NUMBER AND TYPES OF DECISION RULES GENERATED BY JUDGES

Treatment Level	Total Number of Rules	Category					
		Quantitative	Qualitative Clear	Qualitative Vague	Guideline	Comment	Request
Professional	50	16	14	8	1	7	4
Graduate	53	24	18	3	6	1	1
Undergrad.	46	12	19	3	4	8	0
Treatment	85	42	24	9	4	5	1
Control	64	10	27	5	7	11	4
All Judges	149	52	51	14	11	16	5
Treatment Group							
Professional Treatment	28	16	5	4	0	3	0
Graduate Treatment	30	17	6	2	3	1	1
Undergrad. Treatment	27	9	13	3	1	1	0
Professional Control	22	0	9	4	1	4	4
Graduate Control	23	7	12	1	3	0	0
Undergrad Control	19	3	6	0	3	7	0

of experience, undergraduates seemed least capable of extracting quantitative rules when exposed to the treatment conditions.

Rank Order Correlations

A Spearman rank-order correlation coefficient (r_s) was computed for each J between his ranking of the 10 predictors and the order in which they were ranked by the computer model (see Table 8). These values are reported in Table 27.

The data were analyzed using a Two-Way ANOVA and a significant effect was found for the experimental conditions variable ($F=9.35$, $df=1,30$; $p<.01$). That is, Js exposed to the didactic training corresponded more accurately with the computer's ranking of the predictors than did control judges. While no interaction effects were found there was a tendency in the treatment group for correlations to get higher as level of experience decreased. In the control group, this inverse relationship was almost reversed.

The highest correlation coefficient was obtained by a graduate student in the treatment group ($r_s=.88$) while the lowest correlation was reported for an undergraduate student in the control group ($r_s=-.21$).

Variable Weight Correlations

A Pearson product-moment correlation coefficient (r) was computed for each J between the weights he assigned to the 10 predictors and the weights derived for these same variables from the computer model.

TABLE 27

SPEARMAN RANK-ORDER CORRELATIONS BETWEEN JUDGES AND COMPUTER MODEL

Experience Level		Experimental Conditions		Grand Mean
		Treatment	Control	
Professional		.70	.56	
		.30	.34	
		.72	.47	
		.62	.73	
		.55	.54	
		.84	.83	
	Mean	.623	.578	.600
Graduate		.88	-.04	
		.42	.61	
		.85	.58	
		.83	-.07	
		.32	-.09	
		.66	.65	
	Mean	.660	.273	.467
Undergraduate		.81	.01	
		.79	.59	
		.81	.56	
		.58	.76	
		.81	.05	
		.39	-.21	
	Mean	.698	.293	.496
	Grand Mean	.660	.382	.521

The computer model derived weights were determined by a statistical comparison of the Independent F values of the 10 variables. The proportionate magnitude of each variable was determined by the following formula:

$$\text{Weight} = \frac{\text{Independent F Value}}{\text{Sum of Independent F Values}}.$$

For example, the weight for the H DR variable was equivalent to $\frac{14.3889}{148.1972}$ or 9.7. Weights for the other variables were as follows: H SX = 22.9; W SL = 16.4; R-C = 10.2; Husband Stress Factor = 20.3; Wife Stress Factor = 12.1; Husband's age at marriage = 2.1; Wife's age at marriage = 2.7; Number of children = 2.1; Wife's father's occupation = 1.5.

The correlation coefficients are reported in Table 28. The data were analyzed using a Two-Way ANOVA and significant column ($F=4.82$, $df=1,30$; $p<.05$) and interaction ($F=4.80$, $df=2,30$; $p<.025$) effects were found.

The experimental conditions influenced judges to the extent that treatment Js' weights had significantly higher correspondence with the derived weights than did control Js'. The interaction effect was a function of the three levels of experience performing differently depending on the experimental conditions. Under control conditions there was a positive relationship between experience level and magnitude of correlation. However, under treatment conditions, graduates tended to show the highest correspondence ($r=.683$) and professionals the lowest ($r=.457$).

TABLE 28

PEARSON PRODUCT-MOMENT CORRELATIONS BETWEEN JUDGES' WEIGHTS AND COMPUTER MODEL DERIVED WEIGHTS

Experience Level		Experimental Conditions		Grand Mean
		Treatment	Control	
Professional		.70	.74	
		.37	.52	
		.67	.74	
		.00	.84	
		.29	.34	
		.71	.68	
	Mean	.457	.643	.550
Graduate		.93	.21	
		.45	.35	
		.95	.71	
		.66	.25	
		.41	-.10	
		.70	.64	
	Mean	.683	.343	.513
Undergraduate		.89	.29	
		.82	.21	
		.62	.12	
		.66	.89	
		.55	.18	
		.48	-.17	
	Mean	.670	.253	.462
	Grand Mean	.603	.413	.508

The highest correlation coefficient was obtained by a graduate student in the treatment group (r=.95) while the lowest correlation was reported for an undergraduate student in the control group (r=-.17). One J, a professional in the treatment group, assigned the same weight to all 10 variables and, by the nature of the correlation statistic, received an r=.00.

Base Rates

The estimates of population base rates made by each J are reported in Table 29. Although no statistical analyses were conducted on these data, it is apparent that the treatment group came closer to the actual proportions (2:1:1) than did the control group. Undergraduates in the treatment group showed the least discrepancy in estimated vs. real base rates (48:25:27) while their counterparts in the control group reported the greatest discrepancy (30:48:22).

Only four Js were exact in their percentage estimates and three of these were from the treatment group. The hit rates for these Js and the actual number of couples that each categorized into the three criterion groups are presented in Table 30. It is interesting to note that the J with the best hit rate from this special group adhered most closely to the expected base rates while the J with the worst hit rate deviated most dramatically from his reported expectations.

TABLE 29

ESTIMATES OF POPULATION BASE RATES MADE BY JUDGES FOR CRITERION GROUPS

Experience Level	Treatment Conditions			Control Conditions			
	WA	B	AR	WA	B	AR	Totals
	.40	.30	.30	.25	.50	.25	
	.60	.20	.20	.40	.40	.20	
	.35	.25	.40	.55	.40	.05	WA: .41
Professional	.35	.40	.25	.25	.40	.35	B: .34
	.20	.50	.30	.50	.30	.20	AR: .25
	.55	.20	.25	.50	.25	.25	
Total	.41	.31	.28	.41	.37	.22	
	.40	.40	.20	.45	.40	.15	
	.68	.17	.15	.60	.30	.10	
	.50	.25	.25	.30	.50	.20	WA: .46.5
Graduate	.55	.25	.20	.15	.60	.25	B: .33.5
	.55	.25	.20	.20	.50	.30	AR: .20
	.50	.25	.25	.70	.15	.15	
Total	.53	.26	.21	.40	.41	.19	
	.50	.30	.20	.40	.50	.10	
	.60	.15	.25	.15	.45	.40	
Undergraduate	.60	.20	.20	.45	.42	.13	WA: .38.8
	.50	.25	.25	.30	.49	.21	B: .36.3
	.25	.30	.45	.30	.50	.20	AR: .24.9
	.40	.30	.30	.20	.50	.30	
Total	.48	.25	.27	.30	.48	.22	
Totals	.47	.27	.26	.37	.42	.21	

TABLE 30

HIT RATES AND DISTRIBUTION OF CLASSIFICATIONS FOR JUDGES EXACT IN THEIR ESTIMATES OF POPULATION BASE RATES

Characteristics of Judge	Hit Rate	Number of Classifications		
		Well-Adjusted	Borderline	At Risk
Treatment-Graduate	60.0	29	13	18
Treatment-Undergraduate	55.0	17	22	21
Treatment-Graduate	46.7	22	17	21
Control-Professional	43.3	13	23	24
	Expected Frequency	30	15	15

Judges' Self-Evaluations

The mean self-evaluation scores given by Js to reflect their pre- and post-task knowledge regarding the dynamics of college marriage are provided in Table 31 for the different treatment levels and treatment groups. Also listed in this table is the mean increase in knowledge as derived from the two self-evaluations.

Graduate Js were most conservative in appraising their pre-task level of knowledge. As with the MCS data reported earlier in this chapter, undergraduates were most assured of their clinical skills in this area, but not significantly more than the professionals. Undergraduates tended to report the greatest amount of post-task knowledge, as well --especially those in the treatment group.

The increases reported by the individual Js were analyzed using a Two-Way ANOVA and a significant effect was found for the experience level variable ($F=5.50$, $df=2,30$; $p<.01$). Graduate students evaluated the decision-making exercise as being most beneficial in elevating their clinical knowledge. These judges were followed by undergraduates with professionals reporting the least amount of improvement. Amongst the graduate students those in the control group reported the greatest amount of change.

Surprisingly, however, there were no significant differences between the two experimental groups, although treatment Js reported slightly higher increments in

TABLE 31

MEAN SELF-EVALUATION SCORES REFLECTING JUDGES' PRE- AND POST-TASK KNOWLEDGE

Treatment Level	Pre-Task Score	Post-Task Score	Mean Increase
Professional	60.5	65.4	4.9
Graduate	31.1	57.9	26.8
Undergraduate	61.3	71.7	10.4
Treatment	55.3	70.8	15.5
Control	46.6	59.2	12.6
All Judges	51.0	65.0	14.0
Treatment Group			
Professional Treatment	60.2	70.8	10.6
Graduate Treatment	40.8	62.5	21.7
Undergraduate Treatment	65.0	79.2	14.2
Professional Control	60.8	60.0	-0.8
Graduate Control	21.3	53.3	32.0
Undergraduate Control	57.7	64.2	6.5

knowledge. Five Js, four of whom were controls, reported decreases in their post-task scores. Remarks made subsequently by the judges suggest the appropriate interpretation. The negative changes should not imply that the decision-making task was an "unlearning" or "disorienting" experience for these persons. Rather, the decrements probably reflect that the Js' initial evaluations of their marriage-related knowledge were inflated to one degree or another and that the exercise put their "true" knowledge in a more realistic perspective.

CHAPTER VI

DISCUSSION

The Statistical Model: Empirical Relationships in Marital Adjustment

Of the many possible extrapolations which can be made from this research, perhaps as important as any, are the kinds of empirical findings which can be isolated in understanding the relationship between married college life and marital adjustment.

The analyses showed that 18 of the 28 variables used in the statistical model were significantly related to the adjustment criterion. The varying degree of importance of each of the independent measures can be represented qualitatively in terms of one of five ordinal categories. This information is charted in Table 32.

All of the Tharp Discrepancy scores, except for the W II measure, were of unquestionable importance in their contribution to marital adjustment. These findings demonstrate that within-partner role discrepancy increases the likelihood that a marriage is in trouble. The at risk quality of a marriage increases as dissonance between ideal role beliefs and actual role enactments grows for a particular spouse.

TABLE 32

QUALITATIVE IMPORTANCE OF INDEPENDENT VARIABLES USED IN PRESENT STUDY TO MARITAL ADJUSTMENT IN COLLEGE LIFE

Level	Variable
Most Important	H SX H SL W SL Husband Stress Factor
Very Important	Wife Stress Factor W SX R-C H ER H II H DR DIPORE H DIPORE P
Important	W ER W DR DIPORE W Husband's age at marriage Wife's age at marriage Number of children
Questionably Important	W II DAVE P
Not Important	DAVE H DAVE W Length of engagement Length of marriage Number of hours husband employed Total annual income Husband's father's occupation Wife's father's occupation

Also, the more role areas in which dissonance can be found the less probable that the marriage is functioning on a well-adjusted level. In simple words, role consonance creates marital contentment while role dissonance breeds marital disintegration.

A simple architectural analogy is an appropriate one. Each of the role categories can be thought of as a psychological brick which helps hold together the marital structure. Role dissonance can be symbolized as a crack in the respective brick. The greater the dissonance the more evident and expansive the crack. Therefore, as dissonance increases within and over roles, the entire structure becomes more susceptible to stress--it is at risk and capable of toppling.

Of course, there are certain psychological bricks which serve as the foundation of this metaphorical marital structure. If they are strong, other cracks are less dangerous to the total stability; if they are weak, the entire structure is in jeopardy. The present findings suggest that the sexuality and solidarity role categories are important elements in the marital foundation.

As might be expected, deviance from ideal sexual roles is slightly more critical for the husband than for the wife. It is not uncommon knowledge that the American culture is one that sanctions many sex-related "double standards." One of the folklores is that sexual roles, values and behaviors

are more integral to the male partner's experience of his marriage than they are to the female partner. The findings demonstrate that wives, too, place a great deal of value on sexual roles and beliefs and that discrepancies in the enactment of these expectations create conflict for them, as well.

However, from a qualitative perspective, this role area is most important to the husband. A short example should suffice. The man who desires to marry a virgin and later finds out his wife is an experienced woman is more likely to be shaken by this disclosure than the woman who holds a similar conviction only to discover that her husband has been around.

The culturally expected gender-reversal in the area of solidarity was not found. That is, husband and wife both report it is extremely important that their values regarding the cohesiveness of a marriage are met. Husband and wife attach equal and extreme significance to psychological intimacy, emotional sharing, togetherness, understanding and companionship.

In one sense, these findings challenge the traditional, stereotypic and simplistic notion that the formula for a happy marriage is one where the man has his physical needs met and the woman has her emotional needs satisfied. While the characteristics of the sample--a highly educated group of college marrieds--are noticeably different from the

marital population at large, the unusual nature of these findings is still quite impressive. Both men and women bring a cluster of ideal values into marriage regarding both sexuality and emotional solidarity. Dissonance in the enactment of either of these role areas--for either of the two partners--is likely to produce unfulfilled needs and discontent within the marriage.

While the relevance of the intra-spouse role discrepancy scores is unequivocal, the inter-spouse measures derived from the Tharp questionnaire vary in their relationship to the criterion.

Another everyday presumption--that those who have things in common are more likely to have successful marriages--is not substantiated by the DAVE data. Except for the questionable importance of the DAVE P score it was found that discrepancies between partners in the values and attitudes brought into the marriage have no effect on the couple's adjustment. In other words, choosing a mate based on congruity of marriage-related expectations neither increases the potential for success nor failure of the marriage. Likewise, comparative incongruity in partners' values is not portentous in any way.

Once again, this finding may be peculiar to the nature of the sample. It may be that college marrieds are able to transcend the difficulties which often result from differences in values. More importantly, since both partners in

every couple were usually involved in the college community, the spouses already tended to share many similar values and beliefs. Consequently, the possible range of antagonistic values was either considerably decreased or, when dissonance did occur, it took on much less ominous implications.

The DIPORE measures, on the other hand, are between-partner scores capable of discriminating different levels of marital adjustment. The results show that if husband and wife deviate in how they perceive marriage roles being enacted they are more likely to fall into the risk adjustment group. In short, couples that perceive together stay together.

Why this difference between the DAVE and DIPORE variables such that the former has virtually no discriminatory power while the latter can significantly determine level of adjustment? It can be hypothesized that this distinction is due to the fact that the DAVE variable is a "construct" measure, incorporating the terminology of Kelly (1955), and the DIPORE variable is a "communications" measure, in the interpersonal sense of the term.

Lack of similar values and beliefs regarding marriage-roles implies only that the man and woman are operating with different "constructs" about the world of marriage. It does not mean that they are destructive in the manner in which they cope with these differences or even grow from them.

Lack of similar perceptions about how marriage-roles are being enacted implies that the man and woman are experiencing two different worlds of marriage. Their discrepant perceptions are likely to be a function of not communicating about how each experiences what is going on in the marriage and the meaning attached to the behaviors observed. This kind of breakdown in communications can easily be destructive to a marriage.

In this sense, the DAVE measures record the tendency toward "historical" discord between partners while the DIPORE measures focus on the tendency toward "contemporary" discord. Of the two, the data conclusively demonstrate that contemporary discord is the real abrasive factor in a marriage.

The integral importance of monitoring the spouses' relationships with their environment is clearly demonstrated by the discriminating power of the Husband and Wife Stress Factors. The college milieu creates a myriad of daily situations which can either prove rewarding or stressful. The data show that there is an interdependent link between satisfaction derived from the environment and satisfaction within the marriage.

This ecological relationship is not surprising. The findings confirm the assumption that persons experiencing stress in their academic-social interactions are also likely to be troubled in their marriage. This relationship appears

a bit more magnified for the husband. This may be the case since, in many of the couples, only the male partner was enrolled in school and confronting the constant pressures (exams, papers, research projects, etc.) of the academic community.

The R-C score is the other psychological variable found to be of significance to the adjustment criterion. Well-adjusted marriages tend to be those in which the partners express a need for companionship and have that need realized. Unfortunately, the crudeness of the R-C statistic does not allow for a more delicate analysis of the realization-of-companionship dimension. Consequently, a number of questions are still left unanswered. For example, does the couple that reports little need for companionship tend to be better-adjusted than the couple that reports such a need but finds that it goes unfulfilled? Future research related to this area might address such a question.

Only three demographic variables were found to be of importance--although they added very little power to the multivariate model. While ages at marriage and number of children are additional variables which statistically correlate with the criterion, they are rarely crucial in helping to correctly classify any particular couple's adjustment level.

Meehl (1959) suggests the need to jointly utilize life history data with our psychological tests for prediction purposes. Perhaps the multivariate findings suggest that marriage for college-level persons is primarily a psychological phenomenon that is not really biased, shaped or determined by antecedent demographic factors. Or it may be that we have to look for other kinds of life history data rather than the kind gathered here.

Especially striking is the irrelevance of the length of engagement and length of marriage variables--another example of the significance of non-significant test results. While these findings contradict some other research (e.g., Burgess & Cottrell, 1939) they suggest that college couples may be unique in terms of the benefits they derive from lengthy engagements or marriages.

To focus on the issue of engagement, marital success does not appear to be influenced by either short- or long-term engagement periods. However, to emphatically conclude that adjustment is not a function of the length of the period of intimate associations before marriage may be premature. Many of today's college marriages are often the culmination of very intense pre-marital relationships where the couple is likely to have lived together, shared together and made love together--but not under the official guise of a formal engagement.

In short, the traditional notion of engagement may not be as systematically applicable to the college population as it is to the culture at large. In any case, it is an artificial--and perhaps, inaccurate--barometer of the strength and depth of pre-marital attachment. A more appropriate measure of pre-marital attachment needs to be developed.

The Clinical Decision-Making Task

Didactic Implications

For the most part, no striking differences were found between the treatment and control judges. Those allowed to familiarize themselves with the standardization sample before making their predictions were generally no better in their performance than those who were not exposed to the standardization data. On the surface one might conclude that the Didactic Session was inappropriately named since it did not produce any measurable learning differences.

Blumetti (1972) reported a similar finding in his clinical versus actuarial investigation where length of stay in psychotherapy was predicted from MMPI profiles. In fact, Blumetti's small group of pilot judges--who were similar in design to this study's control judges--were much more accurate in their predictions than the regular judges.

The author explains this unexpected difference by suggesting this his didactic judges were asked to

incorporate an overabundance of information which was virtually impossible to process. This "input overload" situation prevented maximum prediction proficiency.

This hypothesis has a great deal of theoretical merit and can easily be applied to the present study. Treatment judges were asked to integrate material about 10 variables over 175 married couples drawn from three groups of marital adjustment. Although most persons seemed to approach this task in a sincere and motivated manner this was a great deal to demand of them. Consequently, the Didactic Session may have generated as much cognitive loss or confusion as it did cognitive gain--to the extent that it produced a "learning stalemate" of sorts.

In other words, the Didactic Session did lead to learning. But, because of the extensive nature of its design, the task produced an equivalent amount of "unlearning," thereby offsetting the possibility of any net learning effect.

In this sense, the teaching value of this sort of decision-making exercise is likely to be camouflaged or underestimated if the results are reviewed in haste. One should not be quick to rule out the treatment methodology used here as a means for facilitating learning about marriage and other areas of mental health. What this study did fail to do was create those ideal conditions which are likely to optimize learning from a decision-making experience.

Tasks such as the one used here have excellent application to numerous training issues. For example, the treatment judges learned more about the relationships between the predictors and criterion, as determined by the computer, than did the control judges. While they were no more accurate in their classifications, treatment judges did show tendencies to be more appropriate in their judgments and more likely to generate decision-rules which were quantitatively based. They also tended to come closer to estimating the actual population base rates.

Skills such as these are worth teaching. Future research, however, must address itself to "cleaning up" the design of the decision-making model. The model presented in this study seems to have its share of faults. Perhaps the number of standardization couples should have been drastically reduced or the number of predictors changed. Good research will hopefully identify those parameters which are most beneficial in producing real learning whose consequences are increased prediction accuracy in human judges and the development of other relevant clinical skills.

Somehow, a mechanism must also be provided which assures that the person will remember (i.e., apply) what he is taught. For example, in this study, it does not appear that the treatment judges made systematic use of the base rates which they had generated. In the few cases where base rate estimates were accurate and the judges used this

information the payoff was noticeable in terms of prediction accuracy. It may well be that simple reinforcement principles can be used to help focus judges on the relevance of attending to such data as base rates, conditional probabilities, etc.

The Role of Experience

The lack of any real differences in accuracy between the three experience groups is somewhat atypical in terms of the decision-making literature. Nonetheless, these results support the hypotheses regarding the role of experience in novel kinds of clinical situations. In general, undergraduates were no less accurate than professionals in determining marital adjustment. Similarly, the kinds of errors made were proportionately the same.

There are a few interpretations which might be suggested. An intriguing one is that a traditional concept such as "clinical experience" may have little meaning in certain situations and may even serve to distort our expectations. The concept is likely to condition the expectation that more clinically experienced judges will naturally have greater prediction skills and will consequently do better on clinical judgment tasks. Such was not the case here.

The decision-making paradigm may well have created an equalitarian learning situation where all judges--irrespective of their level of experience--benefitted from the didactic input. Consequently, any initial differences in prediction

skills were erased by the time the Judgment Session was presented.

However, it is important to note that control judges, too, did not differ on the THR, FPR and FNR measures. This fact would seem to indicate that any initial differences between the three experience groups in terms of prediction skills were purely theoretical (in other words, imaginary).

Blumetti (1972) also reported similar results regarding level of experience but did find significant differences when his judges were compared in terms of an "expertise" variable. The present findings corroborate Blumetti's suggestion that "experience" cannot be equated with "expertise" and that the latter concept probably has more heuristic value. It would appear that decision-making tasks can teach and/or assess expertise and that this is extremely pertinent to the kinds of assumptions made by those individuals who plan graduate training programs.

Certain kinds of differences between the three experience groups can be discovered if one looks closely at some of the other data. As has been demonstrated in other studies, the undergraduates expressed a significantly greater amount of confidence in their judgments. The usual interpretation is that naive, idealistic judges tend to overestimate their facility on certain tasks with which they have never had experience. Therefore, they are very prone to be cocky and tend to exude confidence.

But the opposite interpretation also must be considered. Could it be that clinical experience gradually creates varying degrees of sarcasm, cynicism and self-criticism amongst clinicans so that they eventally lose confidence in their own skills?

There is some reason to believe that the Didactic Session has the greatest influence on the lower levels of experience. For example, as level of experience decreased the Spearman rank-order correlations tended to increase for the treatment judges. The reverse situation was true for the control group.

Similarly, the graduate students in the treatment condition had the highest Pearson product-moment correlations while the professionals reported the lowest. Once again, there seems to be a positive relationship between this correlation variable and experience for the control judges. Also, treatment undergraduates were least discrepant in their base rate estimates.

Finally, it is instructive to look at the self-evaluation data. Here we find that graduate students reported the most improvement in their skills from participating in the decision-making task while professionals expressed the least amount of improvement. Paradoxically, confidence ratings tended to increase the most for professionals in the treatment group.

Perhaps too much can be made of these latter findings. But, in one sense, they suggest an attitudinal or motivational difference between distinct levels of experience. It may well be that the less professionalized person--in this case, the undergraduate and the graduate--is more open, committed and involved with a new learning experience. He enters with fewer preconceptions and predispositions and is more malleable in rearranging his constructs. He is eager to learn and not hesitant to admit ignorance or confusion. If this is so, the unfortunate conclusion may be that experience promotes a kind of mental atrophy in this type of learning set that is likely, over time, to lead to greater apathy toward new learning.

Individual versus Group Decision Making

No consequential findings were reported for the group decision-making task. However, this is a very important research issue which the present study approached from only an explorative point of view. Hopefully, other researchers interested in the dynamics of decision making will choose to pursue this area more vigorously.

The author observed the 12 group decision-making sessions and was quite impressed by the variability in behavior. Some groups were run on an authoritarian basis while others had no leader or coordinator. For some, decision making was based on a simple consensus approach;

for others, judgments were arrived at only after a sharing and processing of information.

The sex variable appeared to play a role in the cohesiveness of a group's interaction. Those groups whose members were of the same sex functioned differently than those composed of both men and women. Also, the management of decision making was noticeably different in those teams made up of active, aggressive individuals when compared with the teams whose members behaved in a less assertive, more passive style.

Of course, other parameters were not even touched upon. For example, group size and its role in fostering or hindering communications must be explored. It would also be informative to design groups whose members were of mixed levels of expertise and experience.

Many of these questions fall within the domain of the social psychologist's research skills. The present study was a very basic attempt to open up these questions to systematic empirical investigation.

Comparison of Statistical and Clinical Approaches

The 55.0 percent cross-validation hit rate for the computer is a fairly impressive statistic considering the use of a three-group criterion. There is reason to believe that this rate might be increased slightly through further restandardization procedures.

The False Negative Rate would appear to be unfortunately high, however. Two out of every five of the At Risk and Borderline couples in the cross-validation sample were identified as being more highly adjusted than they actually were. More specifically, seven of the 15 At Risk couples were given higher adjustment ratings although only two were classified at the well-adjusted level. This kind of error would have detrimental consequences for any kind of intervention program.

On the whole we would have to conclude that the average human judge does no better or worse than the computer when it comes to total number of correct classifications. Nonetheless, 10 individuals in this study had higher THRs than the computer and five other judges tied the 55.0 figure. One judge, a professional in the treatment condition, classified five more correct cases than did the computer.

It would seem that skill in clinical judgment is a characteristic disbributed along a continuum as are many traits or personality variables. Perhaps research can be devoted to identifying what makes judges superior and who these astute judges are in a variety of clinical settings.

There is a slight trend to suggest that the machine excels over man with regard to the False Positive Rate. However, this relationship is reversed when considering the more severe FNR error. Of the 18 treatment judges, all but five beat the machine on this statistic. The Didactic

Session created a relatively conservative perspective amongst the judges such that they were less prone to overestimate marital adjustment than was the computer. This difference is a significant one and a feather in the clinician's hat which should be investigated and understood further.

It is necessary to note that the design for the decision-making task fell short of the "sophisticated clinical" model which Holt (1958) has suggested as the true measure by which to appraise the clinician's skills against the acumen of the statistician. The judges were not presented with any real qualitative material, such as interviews, extensive life histories or projective techniques, to use along with the quantitative data presented on the marriage profiles. As Holt (1970) recently criticized:

> . . . I cannot conceive that any clinical judge could do anything in principle other than apply an informal actuarial method when confronted with nothing but a string of numbers. . . No, if clinical judgment is to be given a chance to show what it can do, it has to have meaningful, qualitative data to work with (though it need not be restricted to such data exclusively). (p. 343)

For this reason, some may argue that the present study shows only that the clinician can function like a machine almost as well as a machine--and sometimes even better than one. Yet, feedback from the judges would imply that they were doing much more than just being mechanical in their predictions.

But the focus should be one of harmony and not of competition. It is only self-defeating to conduct research which emphasizes man versus machine. The more purposeful endeavor is to approach the predictive value of man and machine. The current study shows some areas in which one or the other excels. The goal is to bring together their strengths and minimize their deficiencies.

Community Prevention Proposal

The final purpose of this research is to demonstrate how the data and methodology might be used pragmatically to design, implement and evaluate a community prevention program for college marrieds.

The goals of the program would be threefold:

1. Selection of a risk group within the total married community.
2. Assignment of the "at risk" couples to various forms of intervention services oriented towards understanding the dynamics of the relationship and promoting marital enrichment.
3. Longitudinal follow-up of all couples with especial interest in the marital status of those couples identified in the risk group.

These phases of the proposal will be briefly described below.

Selection Procedures

A "successive hurdles" approach (Meehl & Rosen, 1955) can be used to identify a very select sub-sample of college marrieds. All married couples in attendance within the University community will be given an opportunity to take the same battery of marriage-related tests as used in the present research.

A Locke-Wallace adjustment score and an adjustment rating derived from the statistical prediction formula will be determined for each couple. A couple will be temporarily placed into the risk group if it's MAS score falls below 94 and if it is classified by the computer into the At Risk category.

The Marriage Profile Summaries (MPS) of this potential risk group will be presented to a team of three judges known to be superior in clinical judgment. The judges will be asked to classify the couples into one of the three adjustment categories. Those marriages labeled either Borderline or Well-Adjusted by a majority of the judges will also be eliminated from the risk sample.

Finally, the still remaining couples will be interviewed by a skilled clinician as a qualitative screening device. Those couples who present themselves as being well-adjusted will be eliminated from the final At Risk sample.

This is a rigid screening procedure which will conservatively classify approximately 10-15 percent of the

total population into the At Risk sample. Although this technique will allow for a substantial number of false negatives it will virtually assure the elimination of any false positives from the official risk group. This condition is essential if we are to appropriately evaluate the treatment effects produced by the intervention techniques.

Intervention Techniques

The risk group will be randomly distributed into five sections. Four of these sections will be exposed to various kinds of treatment or secondary prevention services while the fifth section will serve as a control group.

Section A: Role, Communication and Stress (RCS) Therapy Groups

At Risk subjects in this subdivision will be assigned to small groups containing no more than five couples. These groups will meet weekly over a six-month period of time. Each group will have a leader who is experienced in this kind of treatment modality.

The content of the sessions will be specifically designed around the kinds of parameters which the present research has identified as crucial to marital adjustment. The leaders will concentrate on such key issues as marriage roles, communication skills, and the identification of beliefs, values, perceptions and attitudes. Other sessions will be devoted to the relationship between environmental stress and marital harmony.

Various exercises and games will be developed to help the group members focus on these basic issues. While it is difficult to restrict the limits of a group such as this, the leader will be trained to concentrate on a discrete number of dimensions. The structured nature of some of the sessions will assist him in this task.

Section B: Open-Ended Therapy Groups

At Risk subjects in this sub-division will also be assigned to small groups of no more than five couples. Experienced clinicians will function as leaders and the duration will also be six months.

The group leaders will have no special training in the role, communication and stress parameters described above. Rather, they will be instructed to conduct their sessions in an open-ended fashion as they would ordinarily run a marital therapy group, using any techniques or procedures which they judge to be appropriate.

Section C: Role, Communication and Stress (RCS) Individual Therapy

At Risk subjects in this sub-division will be assigned to individual counselors for a six-month program of marital therapy. The therapy approach used will be identical to the techniques used in Section A except for the fact that couples will meet with their therapist on an individual basis instead of using a group format.

Section D: Open-Ended Individual Therapy

At Risk subjects in this sub-division will also be assigned to individual counselors for six months. The counselors will provide the kind of marital therapy they typically conduct in their own practices. No training will be provided these therapists. This is an analogue of Section B using an individual format.

Section E: Control Couples

No treatment services or intervention techniques will be provided.

Longitudinal Follow-up

All married couples who participated in the initial testing will be followed up five years after the original administration of the diagnostic material. During the follow-up phase the test battery will be readministered in all possible cases and other kinds of information will also

be gathered. Of particular interest will be divorce and separation statistics. Other demographic data will be obtained and some new instruments will be developed to measure coping skills within the marriage.

If the prevention proposal presented here is a viable one a number of outcomes are to be expected. While the risk group should generally show a larger incidence of divorce and separation then the rest of the college married sample, those couples exposed to the treatment conditions should report a lower dissolution rate than those control couples exposed to no sort of intervention program. It is predicted that those treatment conditions which emphasize the role, communication and stress parameters (Sections A & C) will produce more successful secondary prevention efforts than the open-ended treatment modalities (Sections B & D).

This theoretical paradigm is represented by the flow-diagram depicted in Figure 3. Hopefully, it exemplifies how the statistical and clinical approaches used in the present research can be incorporated to develop an early warning system for the college community capable of identifying couples who are at risk in marriage and providing them with services that have demonstrated prophylactic value.

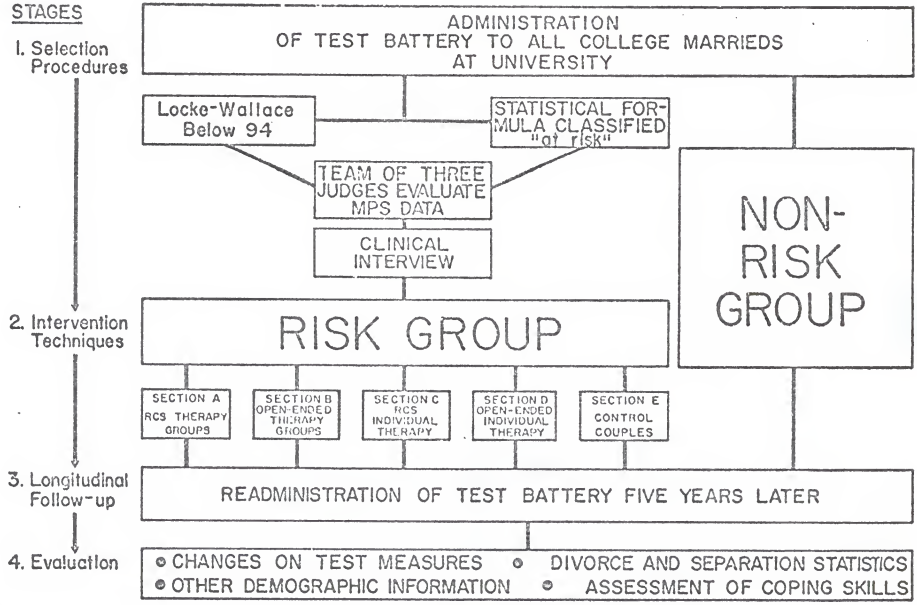

STAGES
1. Selection Procedures
ADMINISTRATION OF TEST BATTERY TO ALL COLLEGE MARRIEDS AT UNIVERSITY
Locke-Wallace Below 94
STATISTICAL FORMULA CLASSIFIED "at risk"
TEAM OF THREE JUDGES EVALUATE MPS DATA
CLINICAL INTERVIEW
NON-RISK GROUP
2. Intervention Techniques
RISK GROUP
SECTION A RCS THERAPY GROUPS
SECTION B OPEN-ENDED THERAPY GROUPS
SECTION C RCS INDIVIDUAL THERAPY
SECTION D OPEN-ENDED INDIVIDUAL THERAPY
SECTION E CONTROL COUPLES
3. Longitudinal Follow-up
READMINISTRATION OF TEST BATTERY FIVE YEARS LATER
4. Evaluation
• CHANGES ON TEST MEASURES
• DIVORCE AND SEPARATION STATISTICS
• OTHER DEMOGRAPHIC INFORMATION
• ASSESSMENT OF COPING SKILLS

Figure 3. Diagrammatic Representation of Community Prevention Proposal

APPENDICES

APPENDIX A

INSTRUMENTS ADMINISTERED TO MARRIED COUPLES

MARRIAGE AND COLLEGE ENVIRONMENT INVENTORY
(Husband Form)

The purpose of this questionnaire is to determine those aspects of the college situation which are most commonly sources of stress or satisfaction to married college students. On the basis of the information gathered we hope to aid the married student. Reports will be presented to several different levels of university administration which will focus attention on ways the university may work more effectively toward improving the living, social, and academic conditions of married student life. Your answers will be added to those of 400 other couples. All of you have been specially selected so that the findings of this study will be representative and of considerable importance. Your cooperation is greatly appreciated.

INSTRUCTIONS:

The items listed relate to aspects of your marriage or the college environment which are experienced as satisfying or as stressful. Consider each item and rate it according to what is generally true. If the aspect generally has been a source of satisfaction or gratification to you, rate it as generally satisfying. If the aspect generally has been a source of stress or dissatisfaction to you, rate it generally stressful.

If you feel that the item has neither been particularly satisfying nor stressful, rate it in the appropriate category -- neither. PLEASE CONSIDER EACH ITEM CAREFULLY, and place in the "neither" category only those items which you feel have not affected you in a satisfying or stressful sense.

Rate as non-applicable those items which have not been part of your life experience. Example: If you have no children, items such as, "Amount of time spent with children," should be rated "non-applicable." However, other items such as, "Amount of encouragement you receive from your wife," should not be rated as "non-applicable." If you receive no encouragement from your wife, consider whether this lack of encouragement has been "generally stressful," "generally satisfying," or has had "neither" effect upon you.

RESPONSE SCALE

Non-applicable	Generally Satisfying	Generally Stressful	Neither
(1)	(2)	(3)	(4)

EXAMPLE:

(06) 1 (2) 3 4 Amount of encouragement you receive from your wife

If you find that this is generally a source of satisfaction to you, you should circle the number 2. The first two digits (06) are for purposes of identification and should be ignored by you.

Please consider each item carefully and circle the appropriate number.

06 1 2 3 4 Associating with other married students
07 1 2 3 4 Monthly earnings relative to cost of living
08 1 2 3 4 Wife's having friends over frequently
09 1 2 3 4 Extent to which we share free time
10 1 2 3 4 Financial expenses related to children

11 1 2 3 4 Working while going to school
12 1 2 3 4 Availability of student financial assistance
13 1 2 3 4 Amount of time you have available for family after studying
14 1 2 3 4 Amount of time your wife has available for family after studying
15 1 2 3 4 Amount of time you take for personal recreation

16 1 2 3 4 Physical fatigue
17 1 2 3 4 Wife's performance of household tasks
18 1 2 3 4 Your assistance with household tasks
19 1 2 3 4 Wife's qualifications for getting good employment
20 1 2 3 4 Being in the "mate-market" atmosphere of the university

21 1 2 3 4 Being with people your own age
22 1 2 3 4 Living in the intellectual and cultural atmosphere of the university
23 1 2 3 4 Learning to budget and manage finances
24 1 2 3 4 Attitude of university toward married students and families
25 1 2 3 4 Amount of time your wife spends on household chores

26 1 2 3 4 Compatibility of schedules
27 1 2 3 4 Having savings to fall back on
28 1 2 3 4 Your arrangements with the draft board
29 1 2 3 4 Borrowing money
30 1 2 3 4 Membership in off-campus organizations

31 1 2 3 4 Membership in on-campus organizations
32 1 2 3 4 Adjusting to a lower income level
33 1 2 3 4 Adjusting to social life and activities of the university
34 1 2 3 4 Attitude of university toward educational interests of wives
35 1 2 3 4 Your wife's not being able to complete her education

36 1 2 3 4 Degree to which you are pursuing your personal interests
37 1 2 3 4 Buying on credit
38 1 2 3 4 Your wife's working full or part-time
39 1 2 3 4 Attending cultural events with your wife
40 1 2 3 4 Your social involvement with your department

41 1 2 3 4 Your social involvement with your wife's department
42 1 2 3 4 Varied schedules from quarter to quarter
43 1 2 3 4 Other's caring for your children while you work or study
44 1 2 3 4 Wide range of attitudes, values, and beliefs on campus
45 1 2 3 4 Requirements of your particular study program

46 1 2 3 4 Requirements of your wife's particular study program
47 1 2 3 4 Your relationship to faculty
48 1 2 3 4 Your wife's relationship to faculty
49 1 2 3 4 Your studying at home
50 1 2 3 4 Your wife's studying at home

51 1 2 3 4 Impersonal nature of the university
52 1 2 3 4 Competition within classes
53 1 2 3 4 Your grades
54 1 2 3 4 Your wife's grades
55 1 2 3 4 Your study habits

56 1 2 3 4 Your wife's study habits
57 1 2 3 4 Studying with children present
58 1 2 3 4 Preparation for quals or comprehensives
59 1 2 3 4 Work on thesis or dissertation
60 1 2 3 4 Effects of mobile situation on children

61 1 2 3 4 Added responsibilities because of children
62 1 2 3 4 Financial assistance from parents
63 1 2 3 4 Your wife's ability to keep pace socially with you
64 1 2 3 4 Extent to which your wife keeps pace intellectually
65 1 2 3 4 Inability to communicate with wife because of terminology or technical nature of your work

66 1 2 3 4 Seeing the end in sight
67 1 2 3 4 Helping friends with their problems
68 1 2 3 4 Wife's assisting with your projects
69 1 2 3 4 Neighbor's reaction to your having or not having children
70 1 2 3 4 Your level of confidence in your ability to be a success in your chosen field

71 1 2 3 4 Wife's level of interest in your major or field of work
72 1 2 3 4 Your present financial status
73 1 2 3 4 Wife's not adhering to her usual schedule
74 1 2 3 4 Degree of appreciation wife shows for your hard work
75 1 2 3 4 Wife's (optimistic/pessimistic) general outlook on life

76 1 2 3 4 Amount of time you arrange for love-making
77 1 2 3 4 Financial independence from parents
78 1 2 3 4 Adjusting to a new community
79 1 2 3 4 Manner in which wife takes care of the children
80 1 2 3 4 Sharing responsibilities for the children

06 1 2 3 4 Amount of companionship with your wife
07 1 2 3 4 Both making sacrifices to complete educational goals
08 1 2 3 4 Initial adjustment to being a student husband
09 1 2 3 4 Being around single students who have less family responsibility
10 1 2 3 4 Your wife being the major or sole support of the family

11 1 2 3 4 Extent to which your wife needs you
12 1 2 3 4 Amount of encouragement you receive from your wife
13 1 2 3 4 The prospect of future financial security
14 1 2 3 4 Postponing things you both want to do
15 1 2 3 4 Sharing responsibilities which are usually the wife's

16 1 2 3 4 Availability of student housing
17 1 2 3 4 Help you received finding housing
18 1 2 3 4 Friendliness and helpfulness of non-university people
19 1 2 3 4 Schools for your children
20 1 2 3 4 Anticipation of how your wife will fill her role in relation to your profession

21 1 2 3 4 Living in a student village
22 1 2 3 4 Living off-campus
23 1 2 3 4 Amount of orientation university gives to married students
24 1 2 3 4 University-provided medical services for wife and family
25 1 2 3 4 Community-provided medical services for wife and family

26 1 2 3 4 Availability of employment for wife
27 1 2 3 4 Unanticipated expenses of a new environment
28 1 2 3 4 Availability of employment for you
29 1 2 3 4 Availability of cultural activities in this city
30 1 2 3 4 Availability of bargain centers and discount stores

31 1 2 3 4 Amount of knowledge concerning this community before coming
32 1 2 3 4 Available recreation outside of university atmosphere
33 1 2 3 4 Finding babysitters
34 1 2 3 4 Privacy from neighbors
35 1 2 3 4 Cost of living in this city

36 1 2 3 4 Playground facilities for children
37 1 2 3 4 Day-care centers
38 1 2 3 4 Knowing that your stay here is temporary
39 1 2 3 4 Availability of playmates for your children
40 1 2 3 4 Adjusting during the first months of marriage

41 1 2 3 4 Wife's views on money matters, and her spending habits
42 1 2 3 4 Wife's interests and hobbies
43 1 2 3 4 Emotional ties with parents
44 1 2 3 4 Maturity of wife in accepting responsibility and handling problems
45 1 2 3 4 Your sex life

46 1 2 3 4 Amount of attention wife gives to the children
47 1 2 3 4 Amount of attention wife gives to you
48 1 2 3 4 Living far away from parents
49 1 2 3 4 Wife's remembrance of sentimental occasions
50 1 2 3 4 Wife's comments on your personal appearance

51 1 2 3 4 Relationship with your in-laws
52 1 2 3 4 Wife's ability to work out differences openly
53 1 2 3 4 Amount of affection shown to you by wife
54 1 2 3 4 Your understanding of how women think, react, etc.
55 1 2 3 4 Way in which wife accepts your shortcomings, failures, etc.

56 1 2 3 4 Amount of knowledge of sex at time of marriage
57 1 2 3 4 Combining adjusting to marriage with going to school
58 1 2 3 4 Wife's display of moods and emotions
59 1 2 3 4 Ability to discuss and communicate with each other
60 1 2 3 4 Openness in discussion of sex interests and concerns

61 1 2 3 4 Wife's habits and idiosyncracies
62 1 2 3 4 Working together on house repair, furnishing, etc.
63 1 2 3 4 Wife's possessive nature
64 1 2 3 4 Way in which problems and arguments are handled in your marriage
65 1 2 3 4 Your wife's health

66 1 2 3 4 Marrying at the time you did
67 1 2 3 4 Differing family backgrounds
68 1 2 3 4 Learning on your own to be effective in marriage
69 1 2 3 4 Your wife's expectations of you as a husband
70 1 2 3 4 Parental attitudes toward your getting married

71 1 2 3 4 Wife's comparing you to other men
72 1 2 3 4 How your expectations for marriage have worked out
73 1 2 3 4 Mutual interests and activities
74 1 2 3 4 Wife's asserting her individuality
75 1 2 3 4 Wife's views on education

76 1 2 3 4 Wife's idea of what your social activities should be
77 1 2 3 4 Having close friends in common
78 1 2 3 4 Manner in which you make decisions together
79 1 2 3 4 Wife's religious beliefs and practices
80 1 2 3 4 Wife's ability as a "handyman"

06 1 2 3 4 Flexibility of wife to change, to accept new ideas
07 1 2 3 4 Wife's need for material goods
08 1 2 3 4 Competition with wife in sports, intellectual activities, etc.
09 1 2 3 4 Wife's political viewpoints
10 1 2 3 4 Strong personality of wife

11 1 2 3 4 Wife's occupational goals
12 1 2 3 4 Effect of your religious beliefs
13 1 2 3 4 Wife's maintaining neat appearance
14 1 2 3 4 Being (or the thought of being) a father
15 1 2 3 4 How decision was made regarding method of contraception

16 1 2 3 4 Your present status of having (or not having) children
17 1 2 3 4 Attitude of wife concerning family planning
18 1 2 3 4 Possibility of an unplanned pregnancy
19 1 2 3 4 The contraceptive being used
20 1 2 3 4 First unplanned pregnancy

21 1 2 3 4 Second or third unplanned pregnancy
22 1 2 3 4 Side effects of contraceptive method
23 1 2 3 4 Religious considerations in planning a family
24 1 2 3 4 Anticipation of your wife's becoming pregnant
25 1 2 3 4 Spacing of children

26 1 2 3 4 Learning on your own how to be a parent
27 1 2 3 4 The first weeks with your new baby
28 1 2 3 4 Your sharing in the care of the baby
29 1 2 3 4 Amount of knowledge of what to expect of children
30 1 2 3 4 Extent to which parenthood has had a maturing effect upon you

Scoring Procedure for Derivation of Stress Factor

$$\text{Stress Factor} = \frac{\#\text{ of Stressful}}{\#\text{ of Satisfying}} \times \frac{\#\text{ of Stressful}}{\#\text{ of Applicable Items}}$$

Hypothetical Examples

	Items	Applicable Items	Stress.	Satis.	Neither	Formula	Stress Factor
Husband A	175	100	80	20	0	$\frac{80}{20} \times \frac{80}{100}$	3.20
Husband B	175	100	20	5	75	$\frac{20}{5} \times \frac{20}{100}$	.80

DERIVATION OF REALIZATION OF COMPANIONSHIP
(R-C) SCORE FROM SERVICES QUESTIONNAIRE

Questions from Which Score Is Derived

Item 83. How adequate is the amount of time and energy you usually have for companionship as a couple?

If Couple Responds	Score
not at all	1
slightly	2
moderately	3
considerably	4
very	5
completely	6

Item 84. Of the various needs you satisfy in marriage, how important to you is companionship?

If Couple Responds	Score
It is the most important need.	3
It is one of the more important needs.	2
It is not one of the more important needs.	1

Scoring Procedure

R-C = Item 83 score X Item 84 score

KELLY-THARP MARRIAGE ROLE QUESTIONNAIRE
(Husband Form)

Your answers to the Questionnaire items will be kept completely confidential. Please do not compare your answers with those of your spouse, since it is important that your responses represent your own personal opinion.

If you have no children, there will be some questions concerning children that you will not be able to answer. Leave these blank. Otherwise, please answer every question.

As you come to each new section, please read carefully the set of response categories that go up the side of the page. They change in meaning from section to section.

Directions: Of the things mentioned below, some are probably essential to a happy marriage, some not desirable, and some not important at all.

Before each statement, draw an X through one of the circles to indicate your opinion of the thing mentioned. What we want is your own personal opinion, whether it agrees with the opinions of other people or not.

YOUR VIEWS ABOUT THE IDEAL MARRIAGE

	Very essential	Usually desirable	Makes little or no difference	Usually not desirable	Decidedly not desirable	How important for the ideal marriage is it:
1.	0	0	0	0	0	That the husband should be the social equal of his wife?
2.	0	0	0	0	0	That the wife should be the social equal of her husband?
3.	0	0	0	0	0	That the husband should be at least equal to his wife in intelligence?
4.	0	0	0	0	0	That the wife should be at least equal to her husband in intelligence?
5.	0	0	0	0	0	That the husband and wife should have similar intellectual interests, such as scientific, literary, musical, etc.?
6.	0	0	0	0	0	That husband and wife should like the same types of amusements, such as cards, dancing, theater, etc.?
7.	0	0	0	0	0	That the husband and wife should engage in the same outdoor sports, such as golf, hiking, swimming, etc.?
8.	0	0	0	0	0	That the husband and wife should each respect the other's religious, political, or ethical convictions and not strive to change them?

YOUR VIEWS ABOUT THE IDEAL MARRIAGE (continued)

9.	0	0	0	0	0	That the wife should be kept fully informed of the family finances and of her husband's business?
10.	0	0	0	0	0	That the father should take an active interest in the discipline and training of the children?
11.	0	0	0	0	0	That the household affairs should be run in a neat, orderly manner?
12.	0	0	0	0	0	That the wife should not have had sexual intercourse with any <u>other</u> man before marriage?
13.	0	0	0	0	0	That the husband should not have had sexual intercourse with any <u>other</u> woman before marriage?
14.	0	0	0	0	0	That after marriage, the wife should be 100% faithful to her husband in regard to sex?
15.	0	0	0	0	0	That after marriage, the husband should be 100% faithful to his wife in regard to sex?
16.	0	0	0	0	0	That husband and wife should be equally fond of social gatherings?
						<u>How important is it to your marriage</u>:
17.	0	0	0	0	0	That you "get ahead" on your job?
18.	0	0	0	0	0	That your home be clean and in order at all times?
19.	0	0	0	0	0	That your wife devote the major part of her interest and energy to her home and family?
20.	0	0	0	0	0	That your home is a place where your family and their friends can relax and enjoy themselves at all times?
21.	0	0	0	0	0	That you and your wife take part in many recreational activities together?

YOUR VIEWS ABOUT THE IDEAL MARRIAGE (continued)

22.	0	0	0	0	0	To have children in your family?
23.	0	0	0	0	0	To have sexual relations every time you desire it?
24.	0	0	0	0	0	That your sexual relations are closely bound up with love and affection?
25.	0	0	0	0	0	That you find pleasure in your sexual relations with your wife?
26.	0	0	0	0	0	That your children are good and well-behaved at all times?
27.	0	0	0	0	0	That your children's ideas and feelings are considered and talked over when family decisions are being made?
28.	0	0	0	0	0	That you, your wife, and your children take part in many recreational activities together?
29.	0	0	0	0	0	That you have sexual intercourse with your wife every time you desire it?
30.	0	0	0	0	0	That your wife find pleasure in her sexual relations with you?
31.	0	0	0	0	0	That your wife be considerate of your feelings about sex?

OPINIONS

This section asks for opinions. There are no right or wrong answers; the best answer to each question is your own personal opinion.

Draw an X through the circle which most closely indicates your feeling about each statement.

OPINIONS (continued)

	I agree very much	I agree a little	Doesn't make any difference	I disagree a little	I disagree very much	
32.	0	0	0	0	0	Women who want to remove the word "obey" from the marriage service don't understand what it means to be a wife.
33.	0	0	0	0	0	Some equality in marriage is a good thing, but by and large, the husband ought to have the main say-so in family affairs.

YOU, YOUR WIFE, AND YOUR MARRIAGE

The Parts You and Your Wife Play: In some way, life is like a play. You each take a turn at playing a number of different parts. At various times, you are a breadwinner, handyman, host, participant in community affairs, friend and companion to your wife, lover and sexual partner to your wife, and father. You have probably found that you are naturally better cast for some of these parts than you are for others. Some men may play the parts of father and breadwinner best. Others may be best fitted for handyman, host, and participant in community affairs. And still others may be best as friends to their wives.

YOU, YOUR WIFE, AND YOUR MARRIAGE (continued)

	Extremely important	Somewhat important	Not at all important	
				How important is it to you that YOUR WIFE should play each of the following parts well?
34.	0	0	0	Housekeeper
35.	0	0	0	Cook
36.	0	0	0	Hostess
37.	0	0	0	Participant in community affairs
38.	0	0	0	Friend and companion to you
39.	0	0	0	Lover and sexual partner to you
40.	0	0	0	Mother
				How important is it to you that YOU should play each of the following parts well?
41.	0	0	0	Breadwinner
42.	0	0	0	Handyman
43.	0	0	0	Host
44.	0	0	0	Participant in community affairs
45.	0	0	0	Friend and companion to wife
46.	0	0	0	Lover and sexual partner to wife
47.	0	0	0	Father

YOU, YOUR WIFE, AND YOUR MARRIAGE (continued)

	Husband much more than wife	Husband somewhat more	Husband and wife about same	Wife somewhat more	Wife much more than husband	In general, who do you think should have more influence in determining the way the family does things in each of the following areas?
48.	0	0	0	0	0	Relationships with relatives
49.	0	0	0	0	0	Choice of friends
50.	0	0	0	0	0	Recreation and social activities
51.	0	0	0	0	0	Earning family income
52.	0	0	0	0	0	Spending family income
53.	0	0	0	0	0	Running the household
54.	0	0	0	0	0	Sexual relations
55.	0	0	0	0	0	Size of family
56.	0	0	0	0	0	Bringing up children

IDEAL FREQUENCY OF RECREATIONAL ACTIVITIES

	Once a day	Twice a week	Once a week	Once a month	Less than once a month	
57.	0	0	0	0	0	How often would you like to have informal get-togethers with other people <u>with</u> your wife?
58.	0	0	0	0	0	How often would you like to have informal get-togethers with other people <u>without</u> your wife?
59.	0	0	0	0	0	How often would you like for you and your wife to play games, chat or watch TV at home without the children or anyone else?
60.	0	0	0	0	0	How often would you like for you and your wife to go out for social or recreational activities without the children or anyone else?
61.	0	0	0	0	0	How often would you like to attend meetings or other activities of groups or organizations <u>without</u> your wife?
62.	0	0	0	0	0	How often would you like to attend such meetings or activities <u>with</u> your wife?
63.	0	0	0	0	0	How often would you like to get together with one or more of the children for fun or recreation <u>at home</u>?
64.	0	0	0	0	0	How often would you like to get together with one or more of the children for fun or recreation <u>away from home</u>?
65.	0	0	0	0	0	How often would you like for all members of the family to get together for some kind of recreation <u>at home</u>?

66. 0 0 0 0 0 How often would you like for all members of the family to get together for some kind of recreation away from home?

	A great deal	Much	Some	A little	None	How much of the housework should usually be done by the following family members?
67.	0	0	0	0	0	Wife
68.	0	0	0	0	0	Husband
69.	0	0	0	0	0	Children
						How much of the physical maintenance of house and yard should usually be done by the following family members?
70.	0	0	0	0	0	Wife
71.	0	0	0	0	0	Husband
72.	0	0	0	0	0	Children

	No children	One (1) child	Two (2) children	Three (3) children	Four (4) or more children	
73.	0	0	0	0	0	When my family is completed, the number of children I prefer is:

	Completely	To a great extent	To some extent	Only slightly	Not at all	
74.	0	0	0	0	0	How well do you feel your wife should understand your ideas and feelings?
75.	0	0	0	0	0	How well do you feel that you should understand your wife's ideas and feelings?

YOUR VIEWS ABOUT YOUR OWN MARRIAGE

						To what extent:
76.	0	0	0	0	0	Are you the social equal of your wife?
77.	0	0	0	0	0	Is your wife your social equal?
78.	0	0	0	0	0	Are you equal to your wife in intelligence?
79.	0	0	0	0	0	Is your wife equal to you in intelligence?
80.	0	0	0	0	0	Do you and your wife have similar intellectual interests, such as scientific, literary, musical, etc.?
81.	0	0	0	0	0	Do you and your wife like the same types of amusements, such as cards, dancing, theater, etc.?
82.	0	0	0	0	0	Do you and your wife engage in the same outdoor sports, such as golf, hiking, swimming, etc.?
83.	0	0	0	0	0	Do you and your wife respect each other's religious, political, or ethical convictions and not strive to change them?
84.	0	0	0	0	0	Do you keep your wife informed of the family finances and of your business?

YOUR VIEWS ABOUT YOUR OWN MARRIAGE (continued)

	Completely	To a great extent	To some extent	Only slightly	Not at all	To what extent:
85.	0	0	0	0	0	Do you take an active interest in the discipline and training of the children?
86.	0	0	0	0	0	Are the household affairs run in a neat, orderly manner?
87.	0	0	0	0	0	Has your wife been faithful to you in regard to sex?
88.	0	0	0	0	0	Have you been faithful to your wife in regard to sex?
89.	0	0	0	0	0	Are you and your wife equally fond of social gatherings?
90.	0	0	0	0	0	Do you "get ahead" on your job?
91.	0	0	0	0	0	Is your home clean and in order at all times?
92.	0	0	0	0	0	Does your wife devote the major part of her interest to her home and family?
93.	0	0	0	0	0	Is your home a place where your family and their friends can relax and enjoy themselves at all times?
94.	0	0	0	0	0	Do you and your wife take part in recreational activities together?
95.	0	0	0	0	0	Do you have sexual relations every time you desire it?
96.	0	0	0	0	0	Are your sexual relations closely bound up with love and affection?
97.	0	0	0	0	0	Have you found pleasure in your sexual relations with your wife during the last three years?

YOUR VIEWS ABOUT YOUR OWN MARRIAGE (continued)

						To what extent:
98.	0	0	0	0	0	Are your children good and well-behaved at all times?
99.	0	0	0	0	0	Are your children's ideas and feelings considered and talked over when family decisions are being made?
100.	0	0	0	0	0	Do you, your wife, and your children take part in many recreational activities together?
101.	0	0	0	0	0	Do you have sexual intercourse with your wife every time you desire it?
102.	0	0	0	0	0	Has your wife found pleasure in her sexual relations with you during the last three years?
103.	0	0	0	0	0	Is your wife considerate of your feelings about sex?
104.	0	0	0	0	0	Do you have the main say-so in family affairs?

YOU, YOUR WIFE, AND YOUR MARRIAGE

	Extremely well	Generally well	Fairly well	Not too well	Not at all well	How well do you think YOUR WIFE plays each of the following parts?
105.	0	0	0	0	0	Housekeeper
106.	0	0	0	0	0	Cook
107.	0	0	0	0	0	Hostess
108.	0	0	0	0	0	Participant in community affairs
109.	0	0	0	0	0	Friend and companion to you

110.	0	0	0	0	0	Lover and sexual partner to you
111.	0	0	0	0	0	Mother

How well do you think YOU play each of the following parts?

112.	0	0	0	0	0	Breadwinner
113.	0	0	0	0	0	Handyman
114.	0	0	0	0	0	Host
115.	0	0	0	0	0	Participant in community affairs
116.	0	0	0	0	0	Friend and companion to wife
117.	0	0	0	0	0	Lover and sexual partner to wife
118.	0	0	0	0	0	Father

	Husband much more than wife	Husband somewhat more	Husband and wife about same	Wife somewhat more	Wife much more than husband	In general, who has more influence in determining the way your family does things in each of the following areas?
119.	0	0	0	0	0	Relationships with relatives
120.	0	0	0	0	0	Choice of friends
121.	0	0	0	0	0	Recreation and social activities

122.	0	0	0	0	0	Earning family income
123.	0	0	0	0	0	Spending family income
124.	0	0	0	0	0	Running the household
125.	0	0	0	0	0	Sexual relations
126.	0	0	0	0	0	Size of family
127.	0	0	0	0	0	Bringing up children

FREQUENCY OF RECREATIONAL ACTIVITIES

	Once a day	Twice a week	Once a week	Once a month	Less than once a month	
128.	0	0	0	0	0	How often do you have informal get-togethers with other people with your wife?
129.	0	0	0	0	0	How often do you have informal get-togethers with other people without your wife?
130.	0	0	0	0	0	How often do you and your wife play games, chat, or watch TV at home without the children or anyone else?
131.	0	0	0	0	0	How often do you and your wife go out for social or recreational activities without the children or anyone else?
132.	0	0	0	0	0	How often do you attend meetings or other activities of groups or organizations without your wife?
133.	0	0	0	0	0	How often do you attend such meetings or activities with your wife?

FREQUENCY OF RECREATIONAL ACTIVITIES (continued)

134.	0	0	0	0	0	How often do you get together with one or more of the children for fun or recreation at home?
135.	0	0	0	0	0	How often do you get together with one or more of the children for fun or recreation away from home?
136.	0	0	0	0	0	How often do all members of the family get together for some kind of recreation at home?
137.	0	0	0	0	0	How often do all members of the family get together for some kind of recreation away from home?

	A great deal	Much	Some	A little	None	
						How much of the housework is usually done by the following family members?
138.	0	0	0	0	0	Wife
139.	0	0	0	0	0	Husband
140.	0	0	0	0	0	Children
						How much of the physical maintenance of house and yard is usually done by the following family members?
141.	0	0	0	0	0	Wife
142.	0	0	0	0	0	Husband
143.	0	0	0	0	0	Children

144. My completed family will probably include this number of children:

- 0 No children
- 0 One (1) child
- 0 Two (2) children
- 0 Three (3) children
- 0 Four (4) or more children

145. How well do you feel your wife understands your ideas and feelings?

146. How well do you feel you understand your wife's ideas and feelings?

- 0 Completely
- 0 To a great extent
- 0 To some extent
- 0 Only slightly
- 0 Not at all

	Completely satisfied	Generally satisfied	About average	Generally dissatisfied	Completely dissatisfied	
147.	0	0	0	0	0	All of us have ideas about what we expect from marriage and what marriage should be like ideally. In terms of the things which you expect from marriage, how satisfied would you say you are with your marriage?

LOCKE-WALLACE MARITAL ADJUSTMENT TEST

1. Check the dot on the scale line below which best describes the degree of happiness, everything considered, of your present marriage. The middle point, "happy," represents the degree of happiness which most people get from marriage, and the scale gradually ranges on one side to those few who are very unhappy in marriage, and on the other, to those few who experience extreme joy or felicity in marriage.

(0)	(2)	(7)	(15)	(20)	(25)	(35)
.	.	.	.	.	.	.
Very Unhappy			Happy			Perfectly Happy

On the following items state the approximate extent of agreement or disagreement between you and your mate. Please circle the appropriate number for your response.

Always Agree	Almost Always	Occasionally Disagree	Frequently Disagree	Almost Always Disagree	Always Disagree
1	2	3	4	5	6

2. Handling family finances.

(5)	(4)	(3)	(2)	(1)	(0)
1	2	3	4	5	6

3. Matters of recreation.

(5)	(4)	(3)	(2)	(1)	(0)
1	2	3	4	5	6

4. Demonstrations of affection.

(8)	(6)	(4)	(2)	(1)	(0)
1	2	3	4	5	6

5. Friends.

(5)	(4)	(3)	(2)	(1)	(0)
1	2	3	4	5	6

6. Sex Relations.

(15)	(12)	(9)	(4)	(1)	(0)
1	2	3	4	5	6

7. Conventionality (right, good, or proper conduct).

(5)	(4)	(3)	(2)	(1)	(0)
1	2	3	4	5	6

8. Philosophy of life.

(5)	(4)	(3)	(2)	(1)	(0)
1	2	3	4	5	6

9. Ways of dealing with in-laws.

(5)	(4)	(3)	(2)	(1)	(0)
1	2	3	4	5	6

10. When disagreements arise, they usually result in: Husband giving in (0), Wife giving in (2), Agreement by mutual give and take (10).

11. Do you and your mate engage in outside interests together? All of them (10) Some of them (8) Very few of them (3) None of them (0)

12. In leisure time do you generally prefer: To be "on the go" ____? To stay at home ____? Does your mate generally prefer: To be "on the go" ____? To stay at home ____?

    scoring: (10) if both at home
    (3) if both on go
    (2) if disagreement

13. Do you ever wish you had not married?
Frequently (0) Occasionally (3) Rarely (8) Never (15)

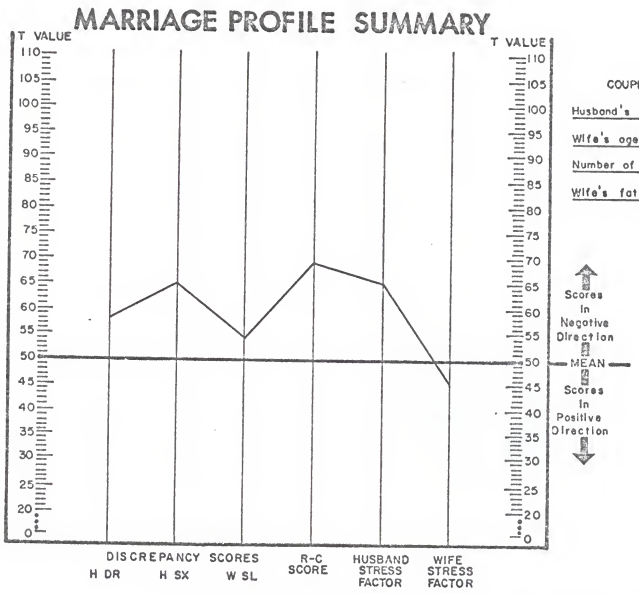
14. If you had your life to live over, do you think you would: Marry the same person (15)? Marry a different person (0)? Not marry at all (1)?

15. Do you confide in your mate?
Almost never (0) Rarely (2) In most things (10)
In everything (10)

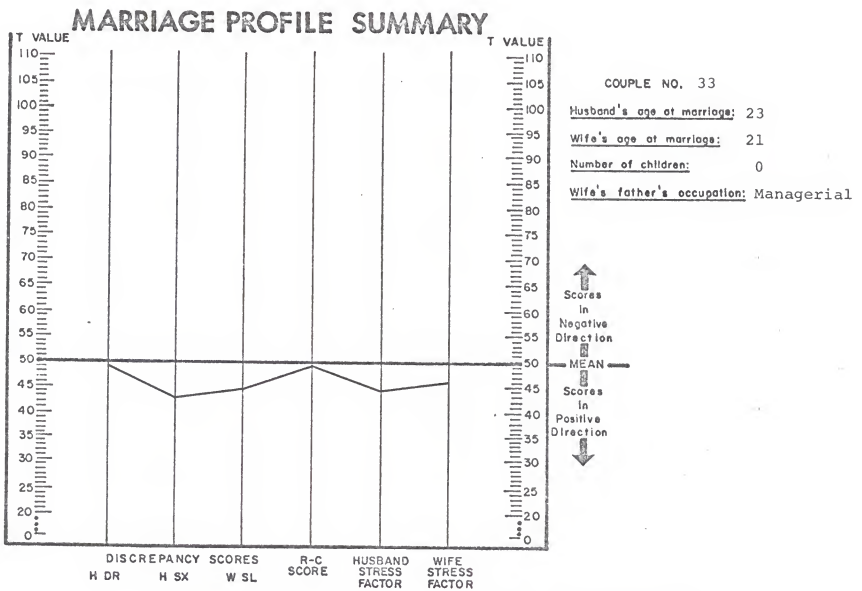
Note: Numbers in parentheses refer to the scoring for that particular response. The scoring procedure for item number 12 is explained above.

APPENDIX B

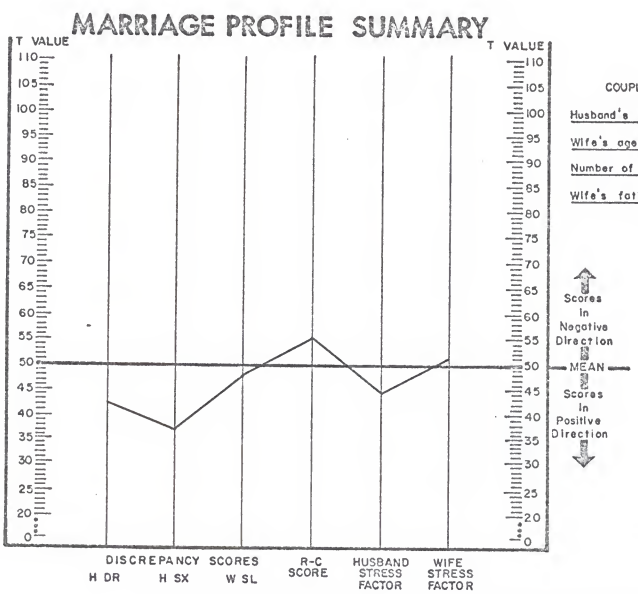
EXAMPLES OF MARRIAGE PROFILE SUMMARY (MPS) FORM TAKEN FROM THE THREE ADJUSTMENT CATEGORIES



Marriage Profile Summary (MPS) of an At Risk Couple
(original material identified with red star)



Marriage Profile Summary (MPS) of a Borderline Couple
(original material identified with yellow star)



Marriage Profile Summary (MPS) of a Well-Adjusted Couple
(original material identified with green star)

APPENDIX C

TEST BOOKLETS USED BY JUDGES

BOOKLET ONE USED BY TREATMENT JUDGES

MARITAL ADJUSTMENT IN A COLLEGE POPULATION: A LEARNING EXERCISE

PHASE I: THE DIDACTIC SESSION

****Please fill out the proper information below before**** commencing with the actual study.

Name_____

Today's Date_____

Academic or Professional Status
(Please check one and provide the appropriate information)

____PhD Clinician
(Years' experience in the field____)
____Grad Student in Clinical
(Years in graduate school____)
____Undergraduate Psychology Major
(Number of psychology courses____)
____Undergraduate Non-psychology Major
(Number of psychology courses____)

YOUR BOOKLET NUMBER IS____

GENERAL INSTRUCTIONS

This study concerns itself with marital adjustment among a population of college marrieds. The purpose of this research is to examine how a decision-making task can be used to "teach" judges how to process test material provided by married couples in order to make predictions about their level of marital adjustment.

Your participation in this study will be divided into two sessions. Today you will participate in Phase I or the Didactic Session. This stage involves more of a learning set than a performance orientation on your behalf. You will be shown Marriage Profiles which contain various psychological and biographical information on a sample of married couples at the University of Florida. These couples have been classified into one of three adjustment groups -- At Risk, Borderline and Well-Adjusted. Your basic task in Phase I will be to analyze the three adjustment groups' profiles so that you develop an understanding of the relationship between the psychological-biographical data and the adjustment criterion.

Tomorrow you will participate in Phase II or the Judgment Session. This stage will involve more of a performance orientation on your behalf. It will allow you to test out the validity of the relationships you have learned through the Didactic Session. You will be shown a new set of Marriage Profiles taken from the same population. These couples also represent one of the three levels of adjustment but their classifications will not be given. Your basic task in Phase II will be to predict the adjustment level for each new couple given the limited information on the Marriage Profiles.

INSTRUCTIONS FOR THE DIDACTIC SESSION

You will now begin Phase I. This exercise is based on the assumption that even if you are completely naive about the dynamics of marital adjustment, careful attention and scrutiny on your part can lead to new knowledge and assessment skills in this area. If you are an attentive and motivated participant you can expect some reasonable payoff in the process. As an added external incentive, a bottle of fine wine will be awarded to the judge in this study who performs most proficiently.

Before proceeding any further estimate how much knowledge you currently possess about the psychological and sociological dynamics of college marriage. Use a 100-point scale. If you have no knowledge in this area you should award one point. If you feel your knowledge is complete in this area you should award 100 points. Assign any other score accordingly.

> I would estimate my knowledge and diagnostic skills in this area with a score of _____.

The Criterion

This session is designed to teach you about the clinical criterion of marital adjustment. While there is some debate as to what is meant by adjustment the couples whose profiles you will view were all evaluated based on an objective instrument derived from the following definition:

> Marital adjustment is the process of adaptation of the husband and wife in such a way as to avoid or resolve conflicts sufficiently so that the mates feel satisfied with the marriage and with each other, develop common interests and activities, and feel that the marriage is fulfilling their expectations.

For conceptual purposes, the adjustment dimension has been demarcated into three sub-categories: At Risk, Borderline and Well-Adjusted. An AT RISK couple is one which -- in relationship to most marriages on this campus -- appears most vulnerable to breakdown and potential dissolution. The AT RISK couple has not necessarily reached the obvious level of marital pathology signified by divorce proceedings. Yet, the psychological momentum of the marriage is heading it in such a direction.

A BORDERLINE couple is one which is not as susceptible to breakdown as those identified as being at risk. However,

the BORDERLINE marriage reflects signs and conditions which would classify it as less healthy than the majority of couples on this campus. Such a marriage occupies a "borderline" status in relationship to the other two adjustment groups -- demonstrating elements of both marital stability and duress.

A WELL-ADJUSTED couple is one which is coping well with the dynamics and demands of married life. Its mates are receiving considerable satisfaction from their partnership. They tend to agree on many issues and when disagreement does arise the discord is resolved democratically and admirably.

The Predictors

You will be presented with Marriage Profiles taken from all three levels of marital adjustment. These profiles contain ten items of information -- six of which are psychological in nature and four of which are biographical in nature. Previous statistical analyses have shown that these 10 variables can be used to predict level of adjustment with a high degree of accuracy for those couples whose profiles you will review. For this reason, these variables can be referred to as predictors of marital adjustment. Of course, some predictors may be more important than others in determining adjustment. Also, a composite arrangement of predictors may lead to a greater number of correct classifications as compared with the individual accuracy of a single predictor.

PSYCHOLOGICAL PREDICTORS:

In all probability, you will have had no clinical or academic experience with the six psychological predictors. These predictors are scores derived from a series of questionnaires which the marriage partners have responded to. They will be described below.

I. Discrepancy Scores. Each spouse brings into the partnership an aggregate of values, beliefs and conventions. These expectations reflect the partner's ideal notions about the manner in which husband and wife should conduct themselves in marriage. Once the partners enact their roles each spouse begins to develop perceptions of the actual or real ways in which husband and wife relate and behave in the marriage setting.

When there is a significant difference between a spouse's expectations and enactments, this state of affairs is referred to as "marriage role dissonance". Similarly, when there is

little difference between a spouse's expectations and enactments, this state of affairs is referred to as "marriage role consonance". The level of dissonance-consonance can be determined by calculating an ideal-real discrepancy score. Discrepancy scores in a negative direction are indicative of role dissonance while discrepancy scores in a positive direction are indicative of role consonance.

Three discrepancy scores are plotted on the Marriage Profile:

1. Husband Division of Responsibility (H DR) -- This score monitors the husband's value system with regard to questions of authority and responsibility for originating, maintaining and terminating role activities within the marriage. Discrepancy scores in a negative direction reflect that his actual perceptions are in conflict with this value system. Discrepancy scores in a positive direction reflect that his expectations are being met with regard to such marital elements as role-sharing, division of influence and masculine authority.

2. Husband Sexuality (H SX) -- This score monitors the husband's value system with regard to questions of premarital chastity, sexual fidelity and sexual gratification. Discrepancy scores in a negative direction reflect that his actual perceptions are in conflict with this sexual value system. Discrepancy scores in a positive direction reflect that his expectations are being met.

3. Wife Solidarity (W SL) -- This score monitors the wife's value system with regard to questions of intimacy, social and emotional integration, togetherness, understanding and companionship. Discrepancy scores in a negative direction reflect that her actual perceptions are in conflict with this solidarity value system. Discrepancy scores in a positive direction reflect that her expectations are being met.

II. Realization of Companionship (R-C). This item is a composite measure of the couple's need for companionship and the amount of time and energy the mates usually have for realizing this need. R-C scores in a positive direction can be interpreted to mean that it is more likely that a couple's need for companionship is great and being realized.

III. Husband and Wife Stress Factor. There are a myriad of aspects in the college situation which are most commonly sources of stress or satisfaction to married college students. This variable monitors the amount of psychological discomfort afforded by everyday situations in the spouse's living

environment. Separate stress factor scores have been determined for both husband and wife. Scores in a positive direction indicate that the mate is generally satisfied with most of his daily activities. Scores in a negative direction indicate that the mate finds the majority of his daily activities to be stress-inducing.

To simplify matters considerably, the psychological predictors have all been standardized and transformed into T scores. If you are unfamiliar with this procedure, in a T distribution the mean is represented by a value of 50 while a value of 10 represents the standard deviation. The six predictor scores for each couple are plotted on the Marriage Profile grid. This grid is similar in design to that of an MMPI profile. Scores which fall below the mean are positive in nature; scores which fall above the mean are negative in nature. The further away from the mean the more atypical the score in comparison with the general standardization group. The standardization group for this study consists of some 235 couples surveyed by the College and Marriage Life Project being conducted on the University of Florida campus. All profiles which you will view have been drawn from this sample.

BIOGRAPHICAL PREDICTORS:

The Marriage Profiles also contain information on four items which are biographical or demographic in nature. These are as follows:

I. Husband's age at marriage. The average age at marriage for the husbands in the standardization group is 22.1 years.

II. Wife's age at marriage. The average age at marriage for the wives in the standardization group is 21.0 years.

III. Number of children. The average number of children reported by couples in the standardization group is 0.85.

IV. Wife's father's occupation. This information serves as a general index of the wife's socioeconomic background. There are five possible occupational categories which reflect increasing levels of socioeconomic status. The highest occupation level is represented by the category PROFESSIONAL. This class includes such occupations as doctor, engineer and teacher. The next highest level is represented by the category MANAGERIAL. Business owner and industrial manager are occupations which exemplify this level. The mid-range occupational category is referred to as CLERICAL/SALES. Jobs which fall into this category are, no doubt, self-

explanatory. The next-to-lowest level is represented by the category SKILLED. Plumbers, electricians, mechanics and the like would qualify under this rubric. The lowest occupational level is represented by the category SEMI-SKILLED. This class includes such occupations as truck driver, waiter and janitor.

The average father of wives in the standardization group maintained an occupation between the MANAGERIAL and CLERICAL/SALES levels, but considerably closer in socioeconomic status to the former.

The above biographical information also appears on the Marriage Profile and is located to the right of the T score grid.

The Task

You should now have some conceptual grasp of the criterion and the predictors which are being used in this study. Perhaps you feel that there are other predictors which relate to marital adjustment and would desire this information on all the couples. Nonetheless, you will be provided with no other information except for the data included on the Marriage Profile.

You will now be presented with three packets of Marriage Profiles drawn from the standardization group. These packets represent the three levels of adjustment described earlier. The profiles of AT RISK couples have been identified with a red star. The profiles of BORDERLINE couples have been identified with a yellow (gold) star. The profiles of WELL-ADJUSTED couples have been identified with a green star.

Your task is to carefully scrutinize the data on as many or few profiles as you care to. From the three groupings attempt to map out in your own cognitive system any relationships which seem to hold between the predictors and the criterion. Use whatever methods possible to define what these relationships are. You may discover that both linear and configural relationships apply. Since each profile is identified with the appropriate marker you may remove profiles from their respective packets in order to make comparisons between groups.

Please be careful in your handling of these profiles since they must be used again by other judges. Do not make any markings on these profiles. Scrap paper will be provided for any notes which you care to take. There is no time limit to this task. Devote as much time to the analysis of the profiles as you feel is needed.

You will be asked to supply some information about the predictor-criterion relationship after you have finished studying the profiles. You may care to glance at the next section of the booklet in order to get an idea of the responses which you will be asked to make.

When you are ready to view the profile packets please signify so and I will present them to you. Proceed to the next section when you feel you are prepared to do so.

EVALUATIONS

I. Now that you have finished studying the three groups of Marriage Profiles, list any rules of thumb, cookbook formulas or if-then statements which you feel have substantial validity and can be used to assign couples into any of the three criterion groups. (An unlikely example would be: "Classify as AT RISK if couple number is odd and number of children is greater than three.") Do not expect your decision rules to be perfectly accurate in making classifications.

1. _____
_____
_____.

2. _____
_____
_____.

3. _____
_____
_____.

4. _____
_____
_____.

5. _____
_____
_____.

II. Rank the ten variables used on the Marriage Profiles in terms of their order of importance to marital adjustment. Then estimate how important these variables are to marital adjustment by assigning each a weight from 0 to 100. You should distribute your weights to represent the relative degree of importance of one variable to another. It is possible for more than one variable to be assigned the same weight. The sum of all weights must total 100.

RANK	VARIABLE	WEIGHT
1	_____	_____
2	_____	_____
3	_____	_____
4	_____	_____
5	_____	_____
6	_____	_____
7	_____	_____
8	_____	_____
9	_____	_____
10	_____	_____
	TOTAL:	_____

III. Assume this sample of profiles is representative of the college married population. What would you estimate are the base rates for At Risk, Borderline and Well-Adjusted conditions in such a population?

CONDITION	BASE RATE
At Risk	_____
Borderline	_____
Well-Adjusted	_____

IV. Please list any other relevant comments about this Didactic Session in the space below.

INSTRUCTIONS FOR TOMORROW'S SESSION

You have now completed Phase I of this study. Tomorrow you will participate in Phase II or the Judgment Session. This session will take place in _____ at _____. You will be participating in this session as both an individual and group judge. So it is extremely important to the design of this study that you not miss tomorrow's session. Until that time, please do not discuss your participation in this project with anyone else, including your fellow judges. Looking forward to meeting with you tomorrow!

BOOKLET TWO USED BY TREATMENT JUDGES

MARITAL ADJUSTMENT IN A COLLEGE POPULATION: A LEARNING EXERCISE

PHASE II: THE JUDGMENT SESSION

****Please fill out the proper information below before**** commencing with the actual study.

Name_____

Today's Date_____

Academic or Professional Status
(Please check one and provide the appropriate information)

____PhD Clinician
(Years' experience in the field____)
____Grad Student in Clinical
(Years in graduate school____)
____Undergraduate Psychology Major
(Number of psychology courses____)
____Undergraduate Non-psychology Major
(Number of psychology courses____)

YOUR BOOKLET NUMBER IS____

INSTRUCTIONS FOR THE JUDGMENT SESSION

The Individual Task

You will now begin Phase II. You will be presented with one packet of 60 Marriage Profiles taken from the same standardization group as yesterday's sample -- college marrieds at the University of Florida.

Your task is to classify each profile into one of the three marital criterion groups -- At Risk, Borderline and Well-Adjusted. To assist you in this decision-making task please make use of any of the information which you provided in Phase I as well as any new insights you may have gained. The "Didactic Booklet" which you used yesterday has been returned to you for this purpose.

For each case classified rate your confidence in that decision on a scale from 50 per cent to 100 per cent. The more certain of your decision the higher the percentage you should mark. If you are positive of your decision, you should mark 100 per cent. If you are only guessing you should mark 50 per cent. Assign your confidence scores in intervals of ten.

Once again, please be careful in your handling of the profiles. Do not make any markings on these profiles. Please make use of the tabulation chart provided in the next few pages in order to score your classifications.

When you are ready to rate the 60 profiles please signify so and I will present them to you. We will proceed to the next task when all judges have completed their individual decisions.

JUDGMENT SESSION: INDIVIDUAL TASK

Tabulation Chart

COUPLE NUMBER	CLASSIFICATION (check one)			CONFIDENCE RATING (check one)					
	At Risk	Borderline	Well-Adjusted	50	60	70	80	90	100
019									
038									
072									
092									
102									
132									
141									
165									
175									
189									
234									
237									
263									
266									
291									
303									
333									
352									
378									
386									

JUDGMENT SESSION: INDIVIDUAL TASK

Tabulation Chart

COUPLE NUMBER	CLASSIFICATION (check one)			CONFIDENCE RATING (check one)					
	At Risk	Borderline	Well-Adjusted	50	60	70	80	90	100
402									
446									
479									
518									
575									
591									
609									
626									
633									
642									
650									
656									
673									
683									
686									
704									
725									
733									
736									
765									

JUDGMENT SESSION: INDIVIDUAL TASK

Tabulation Chart

COUPLE NUMBER	CLASSIFICATION (check one)			CONFIDENCE RATING (check one)					
	At Risk	Borderline	Well-Adjusted	50	60	70	80	90	100
767									
783									
803									
807									
844									
847									
848									
883									
890									
898									
925									
934									
944									
967									
972									
974									
983									
988									
993									
998									

The Group Task

These instructions will be administered verbally when all judges are ready.

FINAL EVALUATIONS

Your participation in this research is all but complete. Yesterday you estimated your pre-exercise level of knowledge about the dynamics of college marriage. How helpful do you feel the Didactic and Judgment Sessions have been in increasing your knowledge base? Use the same 100-point scale to evaluate your post-exercise level of knowledge.

I would now estimate my knowledge and diagnostic skills in this area with a score of _____.

Please include any comments you care to make in the space below. I will be happy to discuss this research with you and to send you an abstract copy of the findings when it is available. Your assistance, time and sincere participation in this project have been greatly appreciated.

Thank you,

Mark B. Lefkowitz

BOOKLET USED BY CONTROL JUDGES

MARITAL ADJUSTMENT IN A COLLEGE POPULATION: A LEARNING EXERCISE

****Please fill out the proper information below before****
commencing with the actual study.

Name_____

Today's Date_____

Academic or Professional Status
(Please check one and provide the appropriate information)

____PhD Clinician
(Years' experience in the field____)
____Grad Student in Clinical
(Years in graduate school____)
____Undergraduate Psychology Major
(Number of psychology courses____)
____Undergraduate Non-psychology Major
(Number of psychology courses____)

YOUR BOOKLET NUMBER IS_____

GENERAL INSTRUCTIONS

This study concerns itself with marital adjustment among a population of college marrieds. The purpose of this research is to examine how judges process test material provided by married couples in order to make predictions about their level of adjustment. This will be accomplished through a decision-making task.

Your participation in this study will be divided into Individual and Group Judgment Sessions. During these two sessions you will be shown Marriage Profiles which contain various psychological and biographical information on a sample of married couples at the University of Florida. These couples represent one of three levels of adjustment -- At Risk, Borderline and Well-Adjusted -- but their classifications will not be given. Your basic task in the Judgment Sessions will be to predict the adjustment level for each couple given the limited information on the Marriage Profiles.

INSTRUCTIONS FOR THE JUDGMENT SESSIONS

This task requires that you be an attentive and motivated participant. As an added external incentive, a bottle of fine wine will be awarded to the judge in this study who performs most proficiently.

Before proceeding any further estimate how much knowledge you currently possess about the psychological and sociological dynamics of college marriage. Use a 100-point scale. If you have no knowledge in this area you should award one point. If you feel your knowledge is complete in this area you should award 100 points. Assign any other score accordingly.

I would estimate my knowledge and diagnostic skills in this area with a score of _____.

The Criterion

You will need some rudimentary information about the clinical criterion of adjustment before proceeding to the actual decision-making task. While there is some debate as to what is meant by adjustment the couples whose profiles you will view were all evaluated based on an objective instrument derived from the following definition:

> Marital adjustment is the process of adaptation of the husband and wife in such a way as to avoid or resolve conflicts sufficiently so that the mates feel satisfied with the marriage and with each other, develop common interests and activities, and feel that the marriage is fulfilling their expectations.

For conceptual purposes, the adjustment dimension has been demarcated into three sub-categories: At Risk, Borderline and Well-Adjusted. An AT RISK couple is one which -- in relationship to most marriages on this campus -- appears most vulnerable to breakdown and potential dissolution. The AT RISK couple has not necessarily reached the obvious level of marital pathology signified by divorce proceedings. Yet, the psychological momentum of the marriage is heading it in such a direction.

A BORDERLINE couple is one which is not as susceptible to breakdown as those identified as being at risk. However, the BORDERLINE marriage reflects signs and conditions which would classify it as less healthy than the majority of couples on this campus. Such a marriage occupies a "borderline" status in relationship to the other two adjustment groups -- demonstrating elements of both marital stability and duress.

A WELL-ADJUSTED couple is one which is coping well with the dynamics and demands of married life. Its mates are receiving considerable satisfaction from their partnership. They tend to agree on many issues and when disagreement does arise the discord is resolved democratically and admirably.

The Predictors

You will be presented with Marriage Profiles taken from all three levels of marital adjustment. These profiles contain ten items of information -- six of which are psychological in nature and four of which are biographical in nature. Previous statistical analyses have shown that these 10 variables can be used to predict level of adjustment with a high degree of accuracy for those couples whose profiles you will review. For this reason, these variables can be referred to as predictors of marital adjustment. Of course, some predictors may be more important than others in determining adjustment. Also, a composite arrangement of predictors may lead to a greater number of correct classifications as compared with the individual accuracy of a single predictor.

PSYCHOLOGICAL PREDICTORS:

In all probability, you will have had no clinical or academic experience with the six psychological predictors. These predictors are scores derived from a series of questionnaires which the marriage partners have responded to. They will be described below.

I. Discrepancy Scores. Each spouse brings into the partnership an aggregate of values, beliefs and conventions. These expectations reflect the partner's ideal notions about the manner in which husband and wife should conduct themselves in marriage. Once the partners enact their roles each spouse begins to develop perceptions of the actual or real ways in which husband and wife relate and behave in the marriage setting.

When there is a significant difference between a spouse's expectations and enactments, this state of affairs is referred to as "marriage role dissonance". Similarly, when there is little difference between a spouse's expectations and enactments, this state of affairs is referred to as "marriage role consonance". The level of dissonance-consonance can be determined by calculating an ideal-real discrepancy score. Discrepancy scores in a negative direction are indicative of role dissonance while discrepancy scores in a positive direction are indicative of role consonance.

Three discrepancy scores are plotted on the Marriage Profile:

1. Husband Division of Responsibility (H DR) -- This score monitors the husband's value system with regard to questions of authority and responsibility for originating, maintaining and terminating role activities within the marriage. Discrepancy scores in a negative direction reflect that his actual perceptions are in conflict with this value system. Discrepancy scores in a positive direction reflect that his expectations are being met with regard to such marital elements as role-sharing, division of influence and masculine authority.

2. Husband Sexuality (H SX) -- This score monitors the husband's value system with regard to questions of premarital chastity, sexual fidelity and sexual gratification. Discrepancy scores in a negative direction reflect that his actual perceptions are in conflict with this sexual value system. Discrepancy scores in a positive direction reflect that his expectations are being met.

3. Wife Solidarity (W SL) -- This score monitors the wife's value system with regard to questions of intimacy, social and emotional integration, togetherness, understanding and companionship. Discrepancy scores in a negative direction reflect that her actual perceptions are in conflict with this solidarity value system. Discrepancy scores in a positive direction reflect that her expectations are being met.

II. Realization of Companionship (R-C) Score. This item is a composite measure of the couple's need for companionship and the amount of time and energy the mates usually have for realizing this need. R-C scores in a positive direction can be interpreted to mean that it is more likely that a couple's need for companionship is great and being realized.

III. Husband and Wife Stress Factor. There are a myriad of aspects in the college situation which are most commonly sources of stress or satisfaction to married college students. This variable monitors the amount of psychological discomfort afforded by everyday situations in the spouse's living environment. Separate stress factor scores have been determined for both husband and wife. Scores in a positive direction indicate that the mate is generally satisfied with most of his daily activities. Scores in a negative direction indicate that the mate finds the majority of his daily activities to be stress-inducing.

To simplify matters considerably, the psychological predictors have all been standardized and transformed into T scores. If you are unfamiliar with this procedure, in a T distribution the mean is represented by a value of 50 while a value of 10 represents the standard deviation. The six predictor scores for each couple are plotted on the Marriage Profile grid. This grid is similar in design to that of an MMPI profile. Scores which fall below the mean are positive in nature; scores which fall above the mean are negative in nature. The further away from the mean the more atypical the score in comparison with the general standardization group. The standardization group for this study consists of some 235 couples surveyed by the College and Marriage Life Project being conducted on the University of Florida campus. All profiles which you will view have been drawn from this sample.

BIOGRAPHICAL PREDICTORS:

The Marriage Profiles also contain information on four items which are biographical or demographic in nature. These are as follows:

I. Husband's age at marriage. The average age at marriage for the husbands in the standardization group is 22.1 years.

II. Wife's age at marriage. The average age at marriage for the wives in the standardization group is 21.0 years.

III. Number of children. The average number of children reported by couples in the standardization group is 0.85.

IV. Wife's father's occupation. This information serves as a general index of the wife's socioeconomic background. There are five possible occupational categories which reflect increasing levels of socioeconomic status. The highest occupation level is represented by the category PROFESSIONAL. This class includes such occupations as doctor, engineer and teacher. The next highest level is represented by the category MANAGERIAL. Business owner and industrial manager are occupations which exemplify this level. The mid-range occupational category is referred to as CLERICAL/SALES. Jobs which fall into this category are, no doubt, self-explanatory. The next-to-lowest level is represented by the category SKILLED. Plumbers, electricians, mechanics and the like would qualify under this rubric. The lowest occupational level is represented by the category SEMI-SKILLED. This class includes such occupations as truck driver, waiter and janitor.

The average father of wives in the standardization group maintained an occupation between the MANAGERIAL and CLERICAL/SALES levels, but considerably closer in socioeconomic status to the former.

The above biographical information also appears on the Marriage Profile and is located to the right of the T score grid.

The Individual Task: Evaluation of the Predictor-Criterion Relationship

You should now have some conceptual grasp of the criterion and the predictors which are being used in this study. Perhaps you feel that there are other predictors which relate to marital adjustment and would desire this information on all the couples. Nonetheless, you will be provided with no other information except for the data included on the Marriage Profile.

You will now begin the Individual Judgment Session. You will be presented with a packet of 60 Marriage Profiles taken

from the larger standardization sample -- college marrieds at the University of Florida.

The first part of this task requires that you carefully scrutinize the data on as many or few profiles as you care to. You will be asked to supply some information about the predictor-criterion relationship after you have finished studying the profiles. Please make use of any of the information provided on the previous pages as well as any clinical insights, hypotheses or hunches you may have about these variables. You may care to glance at the next section of the booklet in order to get an idea of the responses which you will be asked to make.

The Individual Task: Classification of Couples

The second and major part of this task requires that you classify each profile into one of the three marital criterion groups -- At Risk, Borderline and Well-Adjusted. For each case classified rate your confidence in that decision on a scale from 50 per cent to 100 per cent. The more certain of your decision the higher the percentage you should mark. If you are positive of your decision, you should mark 100 per-cent. If you are only guessing you should mark 50 per cent. Assign your confidence scores in intervals of ten. Please make use of the tabulation chart provided in the next few pages in order to score your classifications.

Please be careful in your handling of these profiles since they must be used again by other judges. Do not make any markings on these profiles. Scrap paper will be provided for any notes which you care to take. There is no time limit to this task. Devote as much time to the analysis and classification of the profiles as you feel is needed.

When you are ready to view the 60 profiles please signify so and I will present them to you. Commence your evaluations and ratings when you feel you are prepared to do so. We will proceed to the Group Judgment Session when all judges have completed their individual decisions.

EVALUATIONS

I. Now that you have finished studying the 60 Marriage Profiles, list any rules of thumb, cookbook formulas or if-then statements which you feel may have substantial validity and could be used to assign couples into any of the three criterion groups. (An unlikely example would be: "Classify as AT RISK if couple number is odd and number of children is greater than three.")

1. _____

_____

_____.

2. _____

_____

_____.

3. _____

_____

_____.

4. _____

_____

_____.

5. _____

_____

_____.

II. Rank the ten variables used on the Marriage Profiles in terms of their order of importance to marital adjustment. Then estimate how important these variables are to marital adjustment by assigning each a weight from 0 to 100. You should distribute your weights to represent the relative degree of importance of one variable to another. It is possible for more than one variable to be assigned the same weight. The sum of all weights must total 100.

RANK	VARIABLE	WEIGHT
1	_____	_____
2	_____	_____
3	_____	_____
4	_____	_____
5	_____	_____
6	_____	_____
7	_____	_____
8	_____	_____
9	_____	_____
10	_____	_____
	TOTAL:	_____

III. Assume this sample of profiles is representative of the college married population. What would you estimate are the base rates for At Risk, Borderline and Well-Adjusted conditions in such a population?

CONDITION	BASE RATE
At Risk	_____
Borderline	_____
Well-Adjusted	_____

JUDGMENT SESSION: INDIVIDUAL TASK

Tabulation Chart

COUPLE NUMBER	CLASSIFICATION (check one)			CONFIDENCE RATING (check one)					
	At Risk	Borderline	Well-Adjusted	50	60	70	80	90	100
019									
038									
072									
092									
102									
132									
141									
165									
175									
189									
234									
237									
263									
266									
291									
303									
333									
352									
378									
385									

JUDGMENT SESSION: INDIVIDUAL TASK

Tabulation Chart

COUPLE NUMBER	CLASSIFICATION (check one)			CONFIDENCE RATING (check one)					
	At Risk	Borderline	Well-Adjusted	50	60	70	80	90	100
402									
446									
479									
518									
575									
591									
609									
626									
633									
642									
650									
656									
673									
683									
686									
704									
725									
733									
736									
765									

JUDGMENT SESSION: INDIVIDUAL TASK

Tabulation Chart

COUPLE NUMBER	CLASSIFICATION (check one)			CONFIDENCE RATING (check one)					
	At Risk	Borderline	Well-Adjusted	50	60	70	80	90	100
767									
783									
803									
807									
844									
847									
848									
883									
890									
898									
925									
934									
944									
967									
972									
974									
983									
988									
993									
998									

The Group Task

These instructions will be administered verbally when all judges are ready.

FINAL EVALUATIONS

Your participation in this research is all but complete. Earlier you estimated your pre-exercise level of knowledge about the dynamics of college marriage. How helpful do you feel these Judgment Sessions have been in increasing your knowledge base? Use the same 100-point scale to evaluate your post-exercise level of knowledge.

I would now estimate my knowledge and diagnostic skills in this area with a score of _____.

Please include any comments you care to make in the space below. I will be happy to discuss this research with you and to send you an abstract copy of the findings when it is available. Your assistance, time and sincere participation in this project have been greatly appreciated.

Thank you,

Mark B. Lefkowitz

APPENDIX D

SUMMARY OF NEWMAN-KEULS TESTS

SUMMARY OF NEWMAN-KEULS TEST FOR MEAN DIFFERENCES ON MEAN CONFIDENCE SCORES (MCS): INDIVIDUAL DECISION-MAKING TASK

Differences Between All Judges

	$\bar{X}_P$	$\bar{X}_G$	$\bar{X}_U$
$\bar{X}_P$ = .724	---	.033	.126*
$\bar{X}_G$ = .757		---	.093***
$\bar{X}_U$ = .850			---

*$p < .01$
***$p < .05$

Differences Between Treatment Judges

	$\bar{X}_G$	$\bar{X}_P$	$\bar{X}_U$
$\bar{X}_G$ = .758	---	.008	.103
$\bar{X}_P$ = .766		---	.095
$\bar{X}_U$ = .861			---

Differences Between Control Judges

	$\bar{X}_P$	$\bar{X}_G$	$\bar{X}_U$
$\bar{X}_P$ = .683	---	.072	.156*
$\bar{X}_G$ = .755		---	.084
$\bar{X}_U$ = .839			---

*$p < .01$

SUMMARY OF NEWMAN-KEULS TEST FOR MEAN DIFFERENCES ON MEAN CONFIDENCE SCORE (MCS): GROUP DECISION-MAKING TASK

Differences Between All Groups

	$\bar{X}_P$	$\bar{X}_G$	$\bar{X}_U$
$\bar{X}_P$ = .727	---	.023	.127***
$\bar{X}_G$ = .750		---	.104***
$\bar{X}_U$ = .854			---

***$p < .05$

Differences Between Treatment Groups

	$\bar{X}_P$	$\bar{X}_G$	$\bar{X}_U$
$\bar{X}_P$ = .737	---	.016	.152***
$\bar{X}_G$ = .753		---	.136***
$\bar{X}_U$ = .889			---

***$p < .05$

Differences Between Control Groups

	$\bar{X}_P$	$\bar{X}_G$	$\bar{X}_U$
$\bar{X}_P$ = .718	---	.029	.102
$\bar{X}_G$ = .747		---	.073
$\bar{X}_U$ = .820			---

APPENDIX E

TYPOLOGY OF DECISION RULES GENERATED BY JUDGES

Judge and Rule	Type					
	Quantitative	Qual		Guideline	Comment	Request
		Clear	Vague			
Professional Treatment:						
1. Classify as WA if both stress factors are positive.	X					
2. Classify as AR if H SX negative.	X					
3. If entire profile is positive classify as WA.	X					
4. If 4 psychological variables negative classify AR.	X					
5. If husband greater than 25 at marriage classify as WA.	X					
Professional Treatment:						
1. The lower the mean score of the 6 psychological variables the greater the degree of adjustment.		X				
2. The greater the sum of the husband and wife stress factors the poorer the adjustment.		X				
3. The older the husband at the time of marriage the better the adjustment.		X				
4. Young wives tend to be more poorly adjusted.		X				
Professional Treatment:						
1. Four or more scales greater than 50 T score call high risk. Between three and four scales greater than 50 T score call B. Less than two scales call low risk.	X					
2. The higher the elevation on the scales the higher the risk.		X				
Professional Treatment:						
1. AR is all negative or if either side of profile markedly negative and other side is borderline.	X					
2. AR if marriage borderline and couple quite young when married.			X			

Judge and Rule	Type: Quantitative	Qual: Clear	Qual: Vague	Guideline	Comment	Request
3. B if pattern mixed and no strong points.			X			
4. B if one or two strong negatives and only moderate positives.			X			
5. WA if all positive	X					
6. WA if expectations okay and fairly high stress.			X			
7. Other decision rules that are more complex and more specific.					X	
Professional Treatment:						
1. Classify as WA if all scores are less than 50.	X					
2. Classify as WA if only one score is above 50 unless the wife's father is a professional and they have 2 or more children--then score B.	X					
3. Classify as AR if all scores are above 50.	X					
4. Classify as AR if only one score is below 50 unless the wife's father is skilled or unskilled and they have no children--then score B.	X					
5. Classify as AR if only 2 scores below, at least one of which is a stress factor and the wife's father is professional or managerial. Otherwise classify B for 2 scores only below.	X					
Professional Treatment:						
1. Classify as WA if H DR and H SX are a standard deviation or more below. Classify AR if above one standard deviation.	X					
2. Classify AR if W SL and W Stress Factor are both 1 s.d. or more above mean (more confidence if both, but Wife Stress Factor is pretty good indicator alone).	X					
3. Classify on R-C: WA if 1 s.d. below, AR if 1 s.d. above.	X					

Judge and Rule	Type: Quantitative	Qual: Clear	Qual: Vague	Guideline	Comment	Request
4. Have trouble with biographical. No distribution tendency. May be worth looking at combinations of biographical or biographical and psychological.					X	
5. Seems when Husband Stress Factor and Wife Stress Factor are both close to 1 s.d. above mean or greater, they account for considerable variance on other psychological predictors.					X	
Graduate Treatment:						
1. Classify AR if H SX is greater than 60.	X					
2. Classify AR if either Stress Factor is above 60 or 65.	X					
3. Low R-C (i.e., 35 or lower) usually indicates WA.	X					
4. Given 3 positive discrepancy scores, the R-C score will usually differentiate B from WA.			X			
Graduate Treatment:						
1. Negative H DR is strong predictor of AR. It can be offset by a positive R-C. That is, classify AR if H DR is negative and R-C is not clearly positive. B if H DR is highly negative and R-C is highly positive.		X				
2. H SX--consonance here seems to distinguish B from AR.			X			
3. If Husband's Stress Factor is more negative than Wife's Stress Factor, classify AR.	X					
4. Need to know present economic level. Relationship between that and wife's father's level would seem important.						X
Graduate Treatment:						
1. Initially, ignore biographical data.				X		

Judge and Rule	Type: Quantitative	Qual: Clear	Qual: Vague	Guideline	Comment	Request
2. Classify WA if 0-1 of the psychological predictors over 50; classify AR if 5-6 of the psychological predictors are over 50.	X					
3. Classify B if 2 psychological predictors are over 50 and at least one is over 70. Also classify B if 4 psychological predictors are over 50 but 0 are over 70.	X					
4. In cases where 3 psychological predictors are over 50, look at general pattern of profiles plus biographical data.				X		
Graduate Treatment:						
1. AR if both husband and wife stress over 50.	X					
2. AR if 4 or more over 50.	X					
3. If any one score over 70 then AR.	X					
4. If R-C is over 50 and one stress score elevated classify B.	X					
5. If stress of one mate over 50 and discrepancy score of same mate is elevated relatively classify B.		X				
6. If all 3 Discrepancy scores over 55 classify AR.	X					
7. If more than 2 children don't classify AR.				X		
8. If only H DR or H SX over 50 classify WA.	X					
9. Semi-skilled most likely to be B.		X				
Graduate Treatment:						
1. AR appear to be younger at time of marriage than WA couples.		X				
2. AR couples--majority of factors fall above mean; B generally splits above and below mean; WA generally below mean.	X					
3. WA--H Stress Factor generally below wife's; both close to mean but rarely above for WA; generally opposite for AR couples.		X				
4. Comparative level of H SX, R-C and Husband stress mostly high for AR; W SL and less frequently R-C, low for WA.		X				

Judge and Rule	Type: Quantitative	Qual. Clear	Qual. Vague	Guideline	Comment	Request
5. Wife stress usually lowest factor for high risk; among higher factors for WA--always in vicinity of mean. Less important than husband factors in determining condition of marriage.					X	
Graduate Treatment:						
1. Classify as AR if Husband and/or Wife Stress Factors is over 60.	X					
2. Classify AR if H DR is over 60.	X					
3. Classify AR if any T is over 75.	X					
4. If 4 or more negative scores, regardless of category--AR. If 4 or more positive--classify WA.	X					
Undergraduate Treatment:						
1. The existence of children in a marriage with high stress factors almost guarantees Risk.		X				
2. A highly stressed wife or husband increases Risk (especially if both are).		X				
3. A highly discrepant score in H SX or W SL increases chances of risk, which is guaranteed disaster if combined with a stressed spouse.		X				
4. WA couples show positive R-C scores (though it doesn't seem sign in other couples).		X				
5. One or more extremely negative deviations in any of the predictors (especially Discrepancy and Stress) increases risk.		X				
Undergraduate Treatment:						
1. H DR is not very important in categorizing. The incidence of positive and negative ratings are split between the two.					X	
2. H SX scores are highly negative in AR cases but highly positive in WA cases. Sex expectations must be met for a WA marriage.		X				

Judge and Rule	Type						
	Quantitative	Qual					
		Clear	Vague	Guideline	Comment	Request	
3. In AR cases the W SL variable is negative. Vice versa for WA.		X					
4. RC is not a conclusive variable by which to classify types of marital adjustment.				X			
5. If a majority of predictors are negative then the case is categorized AR.	X						
Undergraduate Treatment:							
1. Couples with a Husband Stress Factors of 50-45 are WA; 51-80 are B; over 80 are AR.	X						
2. Couples with an R-C score of 38-48 are WA; 49-55 are B; over 55 are AR.	X						
3. H-DR--the lower the score the better. WA: 35-42; B: 43-50; AR: over 50.	X						
4. W-SL--WA: 40-50; B: 51-55; AR: over 55.	X						
Undergraduate Treatment:							
1. The higher than average wife's father's status appears most often in WA group.		X					
2. Unless high H DR is counterbalanced the marriage will be AR.			X				
3. Ages of people in WA marriages are generally higher than the mean.		X					
4. Vast majority of the risk marriages are sexually unsatisfactory.		X					
5. Risk can be avoided if stress factors are both substantially lower than mean. This element seems to compensate heavily for discrepancies in other areas.			X				
Undergraduate Treatment:							
1. Any profile showing a T-value of 75 or above always seemed to fall in the AR group.	X						
2. Most cases where the top 3 T values scores deal with only one spouse (all husband or all wife) are the AR group.	X						

Judge and Rule	Type: Quantitative	Qual: Clear	Qual: Vague	Guideline	Comment	Request
3. Most cases where some of the T scores fall below 40 are in the WA group.		X				
4. Most cases where T-values were fairly close to the mean or where a high T-value balanced off with a low T-value leaving an average around the mean were B cases.			X			
Undergraduate Treatment:						
1. If the wife's father's occupation is in the managerial or professional class and if the husband's stress factor is great then classify AR.		X				
2. A couple whose age varies much (10 years) will be in the WA group.	X					
3. If the husband and wife are the same age they will be in the WA group.	X					
4. If the couple has a child and the wife stress factor is high they will be in the B group.		X				
Professional Control:						
1. Negatives for husbands are more predictive of poor adjustment and vice versa.		X				
2. Satisfactions in relationship (H SX and R-C) are needed for good adjustment.		X				
3. Vulnerability greater when wife's father's occupation is professional or managerial and negative H DR.		X				
4. Older ages at marriage and children plus satisfactions equal WA.		X				
Professional Control:						
1. High stress with high discrepancy is okay.			X			
2. High stress with low discrepancy is good.			X			
3. Deduct children from stress factors.				X		
4. Low stress with high discrepancy is deadly.			X			
5. High profile with early marriage is horrible.			X			

Judge and Rule	Type						
	Quantitative	Qual					
		Clear	Vague	Guideline	Comment	Request
Professional Control:							
1. Classify as WA if all T values inside 65 and 35 and biographical indicators not extreme.		X				
2. B if one or two T values greater than 65 and biographical indicators not extreme.		X				
3. Configural--gave added weight for B and AR when more heavily weighted variables occurred together in discrepant (negative) fashion.					X	
Professional Control:							
1. Would like to know husband's major area of study.						X
2. Religious affiliations, if such exist.						X
3. W, DR, W SX, H SL--Apparently these measures are available but not statistically valid.						X
4. If children--conceived before or after marriage?						X
Professional Control:							
1. All of T-scores above 60--higher probability of AR.		X				
2. Age at marriage under 20 for both husband and wife--higher probability of risk.		X				
3. All or most T-scores in positive direction--probability of adjustment.		X				
Professional Control:							
1. Total amount of elevation of the profiles.					X	
2. Extent of elevation of same scales in comparison to others.					X	
3. General stress factors as atmosphere index of current affairs to judge the discrepancy information against.					X	

Judge and Rule	Type: Quantitative	Qual: Clear	Qual: Vague	Guideline	Comment	Request
Graduate Control:						
1. H DR, W SL, R-C over 70 equals AR.	X					
2. Both spouses married at age below 20 consider B.		X				
3. Outlook better if wife in high SES class.		X				
Graduate Control:						
1. Several variables two or more sigmas above T score of 50 should indicate AR marriage.		X				
2. Relationship of one spouse's stress factor to other variables on profile.				X		
3. Relationship of one spouse's stress factor to other's stress factor.				X		
Graduate Control:						
1. Classify as AR those in which most peaks are in the negative direction.		X				
2. Classify as AR especially those where R-C, husband stress factor, wife stress factor and H DR are over 65.	X					
3. Classify as WA those where same variables as indicated in Rule #2 are below 35.	X					
4. Classify as B those profiles which are near the mean.			X			
5. Classify as WA those where most peaks are in positive direction.		X				
Graduate Control:						
1. Classify AR if most or all scores fall toward negative direction.		X				
2. Classify WA if most or all scores fall toward positive direction.		X				
3. Classify B if you have scores falling in both directions with wide discrepancies between plots.		X				
4. Couples over 21-22 with 0-2 children seem to have better chance of adjustment.		X				

Judge and Rule	Type					
	Quantitative	Qual				
		Clear	Vague	Guideline	Comment	Request
Graduate Control:						
1. AR--any couple with 5 scores above mean.	X					
2. Exclude from WA any couple with RC score and one stress factor above mean.				X		
3. WA--any couple with 5 scores below the mean.	X					
Graduate Control:						
1. See if there is substantial age discrepancy. High discrepancy means greater likelihood of strife.		X				
2. Profile with all T scores below 50 gets automatic adjusted score.	X					
3. See if there are multiple stress on the mate. If one mate shows both stress and dissonance likelihood of borderline. Combined with sexual dissonance and lower socioeconomic background likelihood of AR.		X				
4. All T scores above 65 equals AR.	X					
5. Husband and wife discrepancy scores elevated with sexual discrepancy equals B.		X				
Undergraduate Control:						
1. When the husband's age was greater than the wife's at marriage by around 2 years, overall adjustment seemed better.		X				
2. Usually when the wife's father was a professional, the overall marriage picture was better.		X				
3. Number of children has no bearing on determination of criterion group.				X		
Undergraduate Control:						
1. Classify as AR if all scores are negative.	X					
2. Classify as B if both the husband's and wife's stress factors are in the negative range.	X					

Judge and Rule	Type						
	Quantitative	Qual. Clear	Qual. Vague	Guideline	Comment		Request
3. Classify as AR if W SL score and also her stress score are negative.	X						
4. Classify as AR if all of wife's scores are highly negative.		X					
Undergraduate Control:							
1. Wife's solidarity and R-C scores and stress scores weight heavier when children are involved as opposed to a childless marriage.					X		
2. H DR is somewhat less important to the overall picture of marital adjustment.					X		
3. Ages of marital partners have some bearing in adjustment especially if the husband is younger--regards to sex.					X		
Undergraduate Control:							
1. If W SL, H SX, H Stress and W Stress are all negative, then the marriage is a likely risk.		X					
2. In borderline cases, there seemed to be a relationship between H SX and R-C, and W SL and R-C.					X		
Undergraduate Control:							
1. It seems that in a situation where W SL is negative there is a good probability of B.		X					
2. If W SL is negative and R-C is negative with H DR being negative there is a possibility of risk.		X					
3. The fact that more than one child is in the family may have alot to do with maintaining the marriage perhaps though not adjusting to it.					X		
Undergraduate Control:							
1. Examine first R-C score. If R-C is high, it can effectively control stress factors.				X			

Judge and Rule	Type						
	Quantitative	Clear Qual	Vague Qual	Guideline	Comment	Request	
2. Check age--if young possible danger.				X			
3. Even though wife stress is positive, if no children are present uncertainty should be allowed for.					X		
4. In a young man negative H SX coupled with negative stress could encourage adultery, also negative R-C.					X		

REFERENCES

Bernard, Jessie. The distribution of success in marriage. American Journal of Sociology, 1933, 39, 194-203.

Bernard, Jessie. Factors in the distribution of success in marriage. American Journal of Sociology, 1934, 40, 49-60.

Blumetti, A. E. A test of clinical versus actuarial prediction: A consideration of accuracy and cognitive functioning. Doctoral dissertation, University of Florida, 1972.

Bradley, A. D. & Stein, C. Reconceptualization of the clinical and statistical prediction controversy in terms of components of accuracy of interpersonal perception scores. Psychological Reports, 1963, 12, 443-448.

Burgess, E. W. "Factors Determining Success or Failure on Parole," in The workings of the indeterminate sentence law and the parole system in Illinois. Ed. A. A. Bruce. Springfield, Illinois: Illinois State Board of Parole, 1928.

Burgess, E. W. & Cottrell, L. S. Predicting success or failure in marriage. New York: Prentice-Hall, 1939.

Burgess, E. W. & Wallin, P. Engagement and marriage. Philadelphia: J. B. Lippincott Company, 1953.

Chein, I. The logic of prediction: Some observations on Dr. Sarbin's exposition. Psychological Review, 1945, 52, 175-179.

Cline, V. B. & Richards, Jr., J. M. Components of accuracy of interpersonal perception scores and the clinical and statistical prediction controversy. Psychological Record, 1962, 12, 373-381.

Cowen, E. L., Gardner, E. A. & Zax, M. Emergent approaches to mental health problems. New York: Appleton-Century-Crofts, 1967.

Davis, Katherine B. Factors in the sex life of twenty-two hundred women. New York: Harper and Brothers, 1929.

Dahlstrom, W. G. & Welsh, G. S. An MMPI handbook. Minneapolis: University of Minnesota Press, 1960.

Freyd, M. The statistical viewpoint in vocational selection. Journal of Applied Psychology, 1925, 9, 349-356.

Glaser, D. The efficacy of alternative approaches to parole prediction. American Sociological Review, 1955, 20, 283-287.

Glueck, S. & Glueck. Eleanor T. 500 criminal careers. New York: Alfred A. Knopf, Inc., 1930.

Glueck. S. & Glueck, Eleanor T. Unraveling juvenile delinquency. New York: The Commonwealth Fund, 1950.

Glueck, S. & Glueck, Eleanor T. Early detection of future delinquents. Journal of Criminal Law, Criminology and Police Science, 1956, 47, 174-182.

Goldberg, L. R. The effectiveness of clinicians' judgments: The diagnosis of organic brain damage from the Bender-Gestalt test. Journal of Consulting Psychology, 1959, 23, 25-33.

Goldberg, L. R. Simple models or simple processes? Some research on clinical judgment. American Psychologist, 1968, 27, 483-496.

Gough, H. G. "Clinical Versus Statistical Prediction in Psychology," in Psychology in the making. Ed. L. Postman. New York: Alfred A. Knopf, Inc., 1962.

Hamilton, G. V. A research in marriage. New York: Albert and Charles Boni, 1929.

Harris, J. G. Judgmental versus mathematical prediction: An investigation analogy of the clinical versus statistical controversy. Behavioral Science, 1963, 8, 324-335.

Hart, H. Predicting parole success. Journal of the American Institute of Criminal Law and Criminology, 1923, 14, 405-413.

Hart, H. & Shields, W. Happiness in relation to age at marriage. Journal of Social Hygiene, 1926, 12, 403-407.

Harter, A. B. Adjustment of high-school seniors and the marital adjustment of their parents in a Southern California city. Los Angeles: University of Southern California Library 1950.

Holt, R. R. Clinical and statistical prediction: A reformulation and some new data. Journal of Abnormal and Social Psychology, 1958, 56, 1-12.

Holt, R. R. Yet another look at clinical and statistical prediction--Or, is clinical psychology worthwhile? American Psychologist, 1970, 25, 337-349.

Johnston, R. & McNeal, B. F. Statistical versus clinical prediction: Length of neuropsychiatric hospital stay. Journal of Abnormal Psychology, 1967, 72, 335-340.

Kelley, E. L. Concerning the validity of Terman's weights for predicting marital happiness. Psychological Bulletin, 1939, 36, 202-203.

Kelley, E. L. Marital compatibility as related to personality traits of husbands and wives as rated by self and spouse. Journal of Social Psychology, 1941, 13, 193-198.

Kelly, G. A. The psychology of personal constructs. New York: Norton, 1955.

King, C. E. The Burgess-Cottrell method of measuring marital adjustment applied to a non-white southern urban population. Marriage and Family Living, 1952, 14, 280-285.

Kirk, R. E. Experimental design: Procedures for the behavioral sciences. Belmont, California: Brooks-Cole Publishing Co., 1968.

Kirkpatrick, C. Factors in marital adjustment. American Journal of Sociology, 1937(a), 43, 270-283.

Kirkpatrick, C. Community of interest and the measurement of marital adjustment. The Family, 1937(b), 18, 133-137.

Kirkpatrick C. A methodological analysis of feminism in relation to marital adjustment. American Sociological Review, 1939, 4, 325-334.

Landis, J. T. Length of time required to achieve adjustment in marriage. American Sociological Review, 1946, 11, 666-677.

Lang, R. O. "The rating of happiness in marriage." Master's thesis, University of Chicago, 1932.

Lederer, W. J. & Jackson, D. D. The mirages of marriage. New York: W. W. Norton & Company, 1968.

Leslie, G. R. The family in social context. New York: Oxford University Press, 1967.

Lindzey, G. Seer versus sign. Journal of Experimental Research in Personality, 1965, 1, 17-26.

Locke, H. J. Predicting adjustment in marriage: A comparison of a divorced and a happily married group. New York: Henry Holt and Company, 1951.

Locke, H. J. & Karlsson, G. Marital adjustment and prediction in Sweden and the United States. American Sociological Review, 1952, 17, 10-17.

Locke, H. J. & Klausner, W. J. Marital adjustment of divorced persons in subsequent marriages. Sociology and Social Research, 1948, 33, 97-101.

Locke, H. J. & Mackeprang, M. Marital adjustment and the employed wife. American Journal of Sociology, 1949, 54, 536-538.

Locke, H. J. & Wallace, K. M. Short marital adjustment and prediction tests: Their reliability and validity. Marriage and Family Living, 1959, 21, 251-255.

Luborsky, L. Clinicians' judgments on mental health: A proposed scale. Archives of General Psychiatry, 1962, 7, 407-417.

Lundberg, G. A. Case work and the statistical method. Social Forces, 1926, 5, 61-65.

Mannheim, J. & Wilkins, L. T. Prediction methods in relation to Borstal training. London: Her Majesty's Stationery Office, 1955.

Meehl, P. E. Clinical versus statistical prediction. Minneapolis: University of Minnesota, 1954.

Meehl, P. W. Wanted--a good cookbook. American Psychologist, 1956, 11, 263-272.

Meehl, P. E. When shall we use our heads instead of the formula? Journal of Counseling Psychology, 1957, 4, 268-273.

Meehl, P. E. Some ruminations on the validation of clinical procedures. Canadian Journal of Psychology, 1959, 12, 102-128.

Meehl, P. E. & Rosen, A. Antecedent probability and the efficiency of psychometric signs, patterns or cutting scores. Psychological Bulletin, 1955, 52, 194-216.

Merrifield, P. R. & Erickson, E. B. System support for hypothesis generation. Psychological Reports, 1965, 16, 475-490.

Moxley, Ann W. & Satz, P. The effects of statistical information on clinical judgment. Proceedings of the 78th Annual Convention of the American Psychological Association, 1970.

Oskamp, S. The relationship of clinical experience and training methods to several criteria of clinical prediction. Psychological Monographs, 1962, 76 (Whole No. 547).

Phelan, J. G. Rationale employed by clinical psychologists in diagnostic judgment. Journal of Clinical Psychology, 1964, 20, 454-458.

Ray, I. "Statistics of Insanity," in Contributions to mental pathology. Boston: Little, Brown, 1873.

Rogers, L. S., Young, H. H., Cohen, I. H., Dworin, J. & Lipetz, M. E. Marital stability, mental health, and marital satisfaction. Journal of Consulting and Clinical Psychology, 1970, 35, 343-348.

Sarbin, T. R. A contribution to the study of actuarial and individual methods of prediction. American Journal of Sociology, 1942, 48, 593-602.

Satz, P. A block rotation task: The application of multivariate and decision theory analysis for the prediction of organic brain disorder. Psychological Monographs, 1966, 80 (Whole No. 629).

Schroeder, C. W. Divorce in a city of 100,000 population. Chicago: University of Chicago, 1939.

Shagoury, P. "The influence of statistical information on clinical decisions." Master's thesis, University of Florida, 1969.

Shagoury, P. & Satz, P. The effect of statistical information on clinical prediction. Proceedings of the 77th Annual Convention of the American Psychological Association, 1969.

Shneidman, E. S. Suggestions for the delineation of validational studies. Journal of Projective Techniques, 1959, 23, 259-262.

Stricker, G. Actuarial, naive clinical, and sophisticated clinical prediction of pathology from figure drawings. Journal of Consulting Psychology, 1967, 31, 492-494.

Stroup, A. L. Predicting marital success or failure in an urban population. American Sociological Review, 1953, 18, 558-562.

Terman, L. M. Psychological factors in marital happiness. New York: McGraw-Hill, 1938.

Terman, L. M. Prediction data: Predicting marriage failure from test scores. Marriage and Family Living, 1950, 12, 51-54.

Terman, L. M. & Oden, Melita H. The gifted child grows up: Twenty five years' follow up of a superior group. Palo Alto: Stanford University Press, 1947.

Tharp, R. G. Dimensions of marriage roles. Marriage and Family Living, 1963(a), 25, 389-404.

Tharp, R. G. Psychological patterning in marriage. Psychological Bulletin, 1963(b), 60, 97-117.

Tharp, R. G. & Otis, G. D. Toward a theory for therapeutic intervention in families. Journal of Consulting Psychology, 1966 30, 426-434.

Towbin, A. P. When are cookbooks useful? American Psychologist, 1960, 15, 119-123.

Viteles, M. S. The clinical viewpoint in vocational psychology. Journal of Applied Psychology, 1925, 9, 131-138.

Williams, Edith W. Factors associated with adjustment in rural marriages. Ithaca: Cornell University Library, 1938.

Winch, R. F. Personality characteristics of engaged and married couples. American Journal of Sociology, 1941, 46, 686-697.

Wittman, M. P. & Steinberg, L. Follow-up of an objective evaluation of prognosis in dementia praecox and manic-depressive psychoses. Elgin Papers, 1944, 5, 216-227.

BIOGRAPHICAL SKETCH

Mark Barry Lefkowitz was born May 2, 1946, in Brooklyn, New York. Educated in the New York City school system, he graduated from Abraham Lincoln High School in June, 1963. He received the degree of Bachelor of Arts, with a major in psychology, from Brown University in June, 1967. During his junior and senior years he served as Sports Editor of the Brown Daily Herald. In recent years he has been employed as an interviewer for the New York State Employment Service, a vocational counselor for the New York State Division for Youth, and a test consultant for the Stanford Research Institute in conjunction with a Head Start evaluation program in Jacksonville, Florida. He has also worked as a reporter for the Gainesville Sun. In September, 1967, he entered the University of Florida Graduate School as a United States Public Health Service Fellow in the department of psychology. The Master of Arts degree, from the University of Florida, was conferred in June, 1970. During this time he has served as a teaching and research assistant in numerous capacities.

He was appointed Clinical Instructor in the Department of Psychiatry, Ohio State University Medical School, from September, 1971, to August, 1972, and is presently an associate member of the American Psychological Association. Upon completion of the Doctor of Philosophy degree he will serve in an interim position on a specific learning disabilities research project at the University of Florida.

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Paul Satz, Chairman
Professor of Psychology

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Benjamin Barger
Professor of Psychology

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Gerald R. Leslie
Professor of Sociology

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

William C. Mottola
Assistant Professor of Psychology

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Madelaine M. Ramey
Assistant Professor of Psychology

This dissertation was submitted to the Department of Psychology in the College of Arts and Sciences and to the Graduate Council, and was accepted as partial fulfillment of the requirements for the degree of Doctor of Philosophy.

March, 1973

Dean, Graduate School